Combining a land surface model with groundwater model calibration to assess the impacts of groundwater pumping in a mountainous desert basin

Kuai Fang[1], Xinye Ji[1], Chaopeng Shen[*,1], Noel Ludwig[2,†], Peter Godfrey[3†], Tasnuva Mahjabin[1], and Christine Doughty[4]

[1] Civil and Environmental Engineering, Pennsylvania State University, University Park, PA
[2] Region 2 Regional Office, US Department of Agriculture Forest Service, Golden, CO
[3] Arizona State Office, Bureau of Land Management, US Department of Interior, Phoenix, AZ
[4] Earth and Environmental Sciences, Lawrence Berkeley National Laboratory, Berkeley CA

Abstract

The quantification of recharge and trans-valley underflow is needed in arid regions to estimate the impacts of new water withdrawals on the water table. However, for mountainous desert areas, such estimates are highly challenging, due to data scarcity, heterogeneous soils, and long residence times. Conventional assessment employs isolated groundwater models configured with simplified uniform estimates of recharge. Here, we employed a data-constrained surface-subsurface process model to provide an ensemble of spatially distributed recharge and underflow estimates using perturbed parameters. Then, the Model-Independent Parameter Estimation and Uncertainty Quantification (PEST) package was used to calibrate MODFLOW aquifer hydraulic conductivity for this ensemble and reject implausible recharge values. This novel dual-model approach, broadly applicable to mountainous arid regions, was designed to maximally exploit available data sources. It can assimilate groundwater head observations, reject unrealistic parameters, and narrow the range of estimated drawdowns due to pumping. We applied this approach to the Chuckwalla basin in California, USA to determine natural recharge. Simulated recharge concentrates along alluvial fans at the mountain fronts and ephemeral washes where run-off water infiltrates. If an evenly distributed recharge was employed as in conventional studies, it would result in regional biases in estimated drawdown and larger uncertainty bounds. We also note that the speed of groundwater recovery does not guarantee sustainability: heavy pumping induces large hydraulic gradients that initially recover quickly when pumping is halted, but the system may not ultimately recover to pre-pumping levels.

[*] Corresponding author: cshen@engr.psu.edu

[†] Work done at California Desert District, Bureau of Land Management

1. Introduction

For the allocation of groundwater resources in desert or semi-desert areas, the amount of groundwater recharge is an important quantity, since it determines the amount of groundwater that can be renewably extracted. Estimating recharge in mountainous desert basins is especially challenging. Recharge there occurs through spatially sporadic infiltration (Flint et al., 2004) of ephemeral runoff along many washes descending from those mountains (CADWR, 1979) and through associated alluvial fans (Wilson and Guan, 2004). To accurately measure this recharge requires dense, long-term collection of infiltration data in many ephemeral washes and playas. For example, infiltration was found underneath some washes in the Mojave Desert but not others (Izbicki et al., 2000). Such fieldwork and the measurements collected from it are highly valuable and should be strongly supported. These measurements are, however prohibitively expensive and often unavailable. In our study area (the Chuckwalla Basin of southwestern California, to be described in Section 2), there are no data that provide direct measurements of recharge.

Conventionally, groundwater systems have often been modeled with isolated groundwater models such as MODFLOW (Harbaugh, 2005). In that approach, recharge needs to be estimated through independent means, e.g., as a percentage of precipitation (Maxey and Eakin, 1949) or via precipitation-runoff regression (Scanlon, 2004; Wilson and Guan, 2004). Previous environmental impact assessments (EIAs) in the Chuckwalla Basin have used Maxey-Eakin-type recharge estimates, assuming 2 to 10% of precipitation (GEI, 2010; WorleyParsons, 2009). However, this method has strong limitations, because it does not consider the location and mechanism of recharge (Maurer and Berger, 2006). Additionally, using water-balance methods, even small errors in evapotranspiration estimates can result in large-percentage errors in recharge. Large-scale land surface models such as the Community Land Model (Fan et al., 2013) and

PCR-GLOBWB (Sutanudjaja et al., 2017) are also sometimes employed to estimate recharge, but they typically do not have the detailed processes needed to describe desert mountain-front recharge. Physically based integrated hydrologic models, e.g., GSFlow (Markstrom et al., 2008; Tian et al., 2015), HydroGeoSphere (Therrien et al., 2006), ParFlow (Munévar and Mariño, 1999), and PAWS (Shen et al., 2016; Shen and Phanikumar, 2010), calculate recharge as an internal flux. Adapted properly for arid mountainous domains, they can serve as practical tools for recharge estimation. However, the chosen model must provide physically based descriptions of the coupling between surface water run-on infiltration and groundwater. This is because hillslope infiltration and ephemeral-channel percolation are coupled processes controlling the partition of water into soil infiltration, channel percolation, and streamflow yield (Schreiner, McGraw and Vivoni, 2018). Note also that a joint estimation of recharge and the aquifer hydraulic conductivity field, based on groundwater data, is typically not a robust approach, because it has been shown to require a large amount of groundwater head data (Knowling and Werner, 2016) or sufficient prior knowledge (Erdal and Cirpka, 2016).

Integrated hydrologic modeling also faces a range of challenges, including data scarcity in less-inhabited desert areas. First, pedotransfer functions (PTF) (Wösten et al., 2001) are often used to parameterize soil hydrologic functioning in integrated large-scale models, e.g., Maxwell et al., (2015), Mirus, (2015), but desert soil properties differ greatly from what can be inferred from typical PTFs. For example, in desert areas we find closely packed, interlocking rock fragments called “desert pavement” (McFadden et al., 1987) or widely occurring near-surface calcium carbonate formations called “caliche.” These soils are hydraulically distinct from soils elsewhere with similar sand/clay compositions and can vary substantially depending on age (Mirus et al., 2009; Young et al., 2004). Therefore, uncertainty analysis is necessary. Second, recharge that does

occur can take decades to reach the deep water table, requiring expensive long-term simulations on the order of decades or even longer. Finally, aquifer conductivity data are often scarce.

To estimate pumping drawdowns, we need estimates of the regional aquifer conductivity (*K*) or transmissivity (the product of conductivity and thickness). In many desert or semi-desert basins, there are scattered groundwater observations that could be utilized to constrain the highly uncertain *K* field. There might also be some locations with known *K* values estimated from pumping tests or lithologic estimates. In these data-scarce regions, we should use all available information and existing tools to calibrate the spatially heterogeneous *K* field, e.g., using PEST (Doherty, 2003; Tonkin and Doherty, 2009). PEST is a model-independent parameter estimation package that can work with highly parameterized models by exploring covariance structures.

It can be challenging to fully integrate varied sources of information (e.g., groundwater head, soil moisture, and pumping test data) to constrain integrated modeling while also considering the inherent uncertainties. One potential approach is data assimilation (DA), e.g., ensemble Kalman filtering, which updates model states and parameters based on estimating a covariance matrix using ensemble members. DA is a powerful approach that can assimilate multiple data sources, but most applications so far have involved only a few parameters (such as the conductivity value) for a few different zones in a model, e.g., Rasmussen et al. (2015, 2016), Zhang et al. (2015, 2016). By default, DA does not impose geostatistical constraints such as spatial auto-correlation. More recent research has incorporated DA with multi-point geostatistics (MPG) to identify subsurface preferential flow structures (Zovi et al., 2017), but found that the addition of MPG conditioning did not improve their results. Researchers have also employed DA to retrieve the *K* field from synthetic simulations (Pasetto et al., 2015; Xu and Gómez-Hernández, 2016), but the spatial autocorrelation length was assumed and not adjusted. In addition, such specialized

techniques have only been tested in idealized hypothetical scenarios, while real-world cases often possess unique boundary and forcing conditions that are challenging to implement. In real-world cases with scarce *K* data but some hydraulic head observations, it is often not practical to establish usable semi-variograms.

The key research question we pose is, how can we make plausible estimations of recharge ranges in a mountainous desert basin, given scattered observations of soil moisture, groundwater head, and limited or uncertain hydrogeology knowledge? Owing to limited observations, we did not attempt to estimate a probability distribution of the recharge, but rather to put a bound on the estimate. Furthermore, given uncertain parameters, we want to know how the conditioning of the models on available information influences the estimated pumping drawdown,. In this paper, we demonstrate the usefulness of a dual-model approach to best exploit multiple sources of data for the purpose of putting bounds on natural recharge in the basin and constraining future projections of pumping effects. It is possible that further research in DA will better incorporate different forms of geostatistical constraints. However, here we take the simpler approach of coupling a process-based model with a widely used parameter estimation package that can impose geospatial autocorrelation. Importantly, the straightforward nature of the framework makes it more accessible to practitioners, who may also wish to replace the hydrologic model and calibration procedures with ones of their own choosing.

2. Sites and Methods

We employed a data-constrained integrated hydrologic model to provide recharge estimates for calibrating a groundwater model. The sparsely inhabited study region poses significant challenges for modeling, especially in terms of characterizing hydrogeology and soils. Groundwater and geologic data have not been systematically collected. However, scattered

information can be found across multiple reports from resource reconnaissance, mining operations, and environmental impact analysis. To assemble the model, we needed to compile and integrate existing quantitative and qualitative knowledge, as well as make some assumptions (to be described in Section 2.3.2), including a flat top for the one of the aquifer formations, as described below. Given parameter uncertainties, we ran an ensemble of simulations using a process-based integrated hydrologic model, PAWS+CLM (Section 2.3.1). Recharge estimates were extracted from these simulations (denoted by simulations #0, #1, etc.) and were subsequently fed into MODFLOW+PEST, a parameter calibration package capable of estimating spatially distributed parameters. The recharge values that could not adequately describe groundwater observations were rejected, while those retained were used to make predictions (Figure 1). Only the retained recharge and parameter sets were used to estimate the impacts of pumping. In the following, we first describe physiographic features of the basin and our field measurements. Then we describe how the model was assembled from raw information.

2.1. Basin physiographic properties

We examine the Chuckwalla Basin (Figure 2a, 6,712 km^2), which is located west of the city of Blythe near the Colorado River in California, generally within the Mojave and Sonoran Deserts. Pumping activities have been proposed for new water users in this basin, including solar energy projects and a pumped energy storage project (Table 1). The approved operations will collectively extract 2.3 $\times 10^6$ $m^3 yr^{-1}$ from local aquifers on an ongoing basis. Additionally, the Eagle Mountain Pumped Storage project (EMPS), is permitted to extract almost 1×10^7 $m^3 yr^{-1}$ during the 4-yr initial-fill phase (FERC, 2012). We are interested in predicting the impacts of such pumping on the local aquifer.

147 The basin has a hot desert climate, with average January and July temperatures of 4°C and 43°C,
148 respectively, and an 18-year annual average rainfall of 95 mm. There are no perennial water
149 bodies within the basin. About 30% of the basin is mountainous terrain rising abruptly from the
150 valley floor. Generally, the basin floor slopes gently downward from northwest to southeast.
151 Physiographically, there is an upper (western) and lower (eastern) portion of the Chuckwalla
152 Valley proper, with a subtle surface water divide between the Palen (western) and Ford Dry Lake
153 (eastern) playas. Some groundwater underflow enters from the Pinto Valley Basin in the
154 northwest. The smaller Orocopia Valley Basin (to the west or southwest of Desert Center) likely
155 contributes negligible underflow. This is otherwise a closed basin, except along its eastern
156 boundary with the Palo Verde Valley and the Colorado River. Metamorphic and igneous bedrock
157 composing the surrounding mountains is assumed to be impervious (WorleyParsons, 2009).

158 The basin has complex soil configurations, yet widely used soil surveys have poor coverage for
159 this region. SSURGO, the most detailed soils survey available nationally, contains only one soil
160 type for most of the Chuckwalla Valley and mountains, with no descriptions of soil thickness or
161 water retention properties. Generally, the mountains have considerable exposed bedrock with
162 thin, sandy soils and washes or alluvial drainages supplying sediment to the valley floor. No data
163 exist to describe the thickness of these soils, but our field reconnaissance suggested an average of
164 30 cm. In contrast, the valley is mantled by much thicker soils and unconsolidated materials such
165 as loamy sand alluvium with interlacing desert pavement (Figure 3a), coarse, steep alluvial fans
166 at the mountain feet, where brush can grow (Figure 3b), and clay-rich playas near the center
167 (USGS, 1995). A satellite image of the upper basin, with visible traits of the washes, is provided
168 in Figure 3c.

Well logs indicate that the alluvium layer in the valley (interbedded sands and gravels with discontinuous clay) varies between 210 m and 366 m in thickness (CADWR, 1979). The water table can often be found in this layer. Depth to water table ranges from 150 m near Desert Center to 6.4 m near Palen Dry Lake, where groundwater may emerge and slowly evaporate. With respect to the geology beneath the alluvium and above the bedrock, the valley can be roughly divided into two separate zones. These are the water producing Bouse Formation in the lower valley (Metzger et al., 1973), and a fine grained lacustrine silt/clay layer underlying the upper valley that acts as an aquitard. The Bouse Formation is a Pliocene marine and estuarine sequence composed of limestone, clay, silt, and sand (Owen-Joyce et al., 2000). Well logs suggest its surface is flat (Stone, 2006; WorleyParsons, 2009). However, the location of the western boundary of Bouse is not entirely clear. It is not noted west of Desert Center (GEI, 2010). A Miocene Fanglomerate aquifer unconformably underlies the Bouse, but their interface is indistinct. Shrubs and other specialized desert plants are most abundant on the valley floor, associated with alluvial fans and washes (Figure 3b shows a picture at the foot of an alluvial fan). The productive aquifers are the unconfined alluvium in the upper basin and the Bouse Formation in the lower basin. The use of this information in the model is described in Section 2.3.2.

2.2. In situ measurements

Besides the five regular meteorological stations in the basin, two new stations have been installed recently with soil moisture probes. These include two Soil Climate Analysis Network (SCAN) stations near Desert Center and Ford Dry Lake (Figure 2a). Soil moisture data have been collected at depths of 5, 10, 20, and 50 cm below ground surface (bgs) at the SCAN stations since late 2011. A monitoring well, CWV1, was completed in 2012 to 300 m bgs, near the groundwater outflow of the basin, to collect groundwater and geophysical data in separate

192 aquifer intervals, including natural gamma, electric resistivity, and sonic logs (Everett, 2013).
193 Using a linear sonic transit time formulation corrected by gamma-log-based clay fraction data
194 (RMC, 1990), investigators calculated porosity at different depths of the well.

195 Well records from USGS Groundwater Watch, California Department of Water Resources
196 (CADWR), and well logs in EIA reports were compiled for calibrating the groundwater flow
197 model. Some of these wells have estimates of transmissivity and conductivity derived from
198 pumping tests, which were also utilized.

199 **2.3. Surface-subsurface processes modeling**

200 *2.3.1 PAWS+CLM*

201 The Process-based Adaptive Watershed Simulator coupled to the Community Land Model
202 (PAWS+CLM) is a comprehensive and computationally efficient model representing the
203 whole-land phase of the hydrologic cycle (Ji and Shen, 2018; Niu et al., 2017, 2014, Shen et al.,
204 2016, 2014, 2013; Shen and Phanikumar, 2010) and reactive transport (Niu and Phanikumar,
205 2015) on a rectangular grid (Figure 4a). CLM4.0 is a comprehensive community-shared land
206 surface process model describing water, carbon, nitrogen, energy, and other cycles (Oleson et al.,
207 2010). In our coupling, CLM predicts the surface energy, vapor, and carbon transfers between
208 land and atmosphere. It solves soil temperature, snow freeze/thaw, condensation, sublimation,
209 photosynthesis, carbon allocation, soil carbon decomposition, and nitrogen cycling. These
210 biogeophysical and biogeochemical processes provide boundary conditions for the hydrologic
211 calculations of the PAWS modules (Figure 4b) discussed in the next paragraph. For this case, we
212 used CLM4.0, but adapted a micro-topographic surface ponding parameterization from CLM4.5
213 as described later in this section. The PAWS+CLM model provides recharge estimates that can be
214 extracted and provided to the groundwater model calibration package MODFLOW-PEST, as

discussed in Section 2.4. MODFLOW-PEST can then provide calibrated K fields back to PAWS+CLM. However, for this study, this feedback was not employed.

Hydrologic processes including soil water, groundwater, surface water, and multi-way exchanges are provided by PAWS. The soil hydrology module in CLM was replaced with its PAWS counterpart (Shen et al., 2014, 2013), and the CLM soil water storage was deactivated. PAWS solves 2D overland flow, channel flow, vertical soil water flow, and quasi-3D saturated groundwater flow. The overland flow is governed by the diffusive wave equation. Soil water in the unsaturated zone is governed by the 1D Richards equation, while groundwater is governed by the saturated confined or unconfined flow equation. PAWS also describes dynamical, multi-way exchanges between surface water, soil, groundwater, and channel compartments. Mass conservation is enforced. PAWS has been verified by comparison with analytical solutions and other full-3D models (Maxwell et al., 2014). In addition, PAWS+CLM, deployable globally, has satisfactorily reproduced a wide variety of field observations, including streamflow, groundwater depths, leaf area index, evapotranspiration, soil moisture and temperature, and water storage (Ji et al., 2015; Pau et al., 2016; Riley and Shen, 2014).

Given that the model has been described in previous papers, here we highlight the processes relevant to desert recharge and the specific questions posed. We have accounted for three possible recharge sources from PAWS+CLM to be passed to MODFLOW+PEST: (1) run-on infiltration (also called leakance) from the overland flow in the washes (Q_{og}), (2) direct soil column recharge (R), and (3) mountain-front subsurface flow (Q_{Msub}). Another theoretically possible recharge mechanism is the leakage from the Colorado River, but because of the groundwater boundary configuration (Section 2.4.1) and the position of the river in the most downstream part of the domain, it has little influence on the simulated recharge in the valley.

To determine leakance, surface water is divided into flow and ponding domains, which are two conceptualized storage compartments in the same grid cell. The ponding domain exchanges vertically with the soil column, while the flow domain flows laterally between grid cells. The ponding domain contributes runoff to the flow domain. On the other hand, if the flow domain has enough water depth, water can propagate into the ponding domain and infiltrate through the soil matrix. The flow direction and velocity of the flow domain are determined automatically as governed by the 2D diffusive wave equation. The flow domain is conceptualized to concentrate in the lowest part of the grid cell, in a fraction of the cell termed f_w**.** Following a micro-topographic parameterization in CLM4.5 (Oleson et al., 2013), f_w is estimated as:

$$f_w = \frac{1}{2}\left(1 + erf\left(\frac{h_f}{\sigma_m\sqrt{2}}\right)\right)$$

$$\sigma_m = \left(\beta + \sigma_{max}^{1/\eta}\right)^{\eta}$$

where h_f is the flow domain water depth, β is the slope, σ_{max} and η are parameters. The flow domain may exchange with the groundwater using a Darcy-type leakance concept (Gunduz and Aral, 2005), i.e., $q_b = K_b \frac{h_f + D}{\min(2,D)}$ when $h_f > 0$, where K_b is the leakance conductivity, and *D* is the depth to the water table. In most of the Chuckwalla Valley, because the groundwater table is very deep, this formulation amounts to a vertical hydraulic gradient of one (free drainage). K_b is tied to the saturated conductivity of the soil and thus the leakance is adjusted during the calibration. At the same time, evaporation of the flow-domain water occurs at a potential rate (estimated by the Penman Monteith equation) scaled by f_w. Evaporation takes place before leakance in the model. Thus, if evaporation is strong enough, there will be no leakance.

Water flow in the unsaturated soil column is modeled by the 1D Richards Equation, which incorporates the nonlinear response of soil pore pressure and unsaturated conductivity to changes in water content. Underneath the phreatic water table, soil water interacts with the saturated groundwater, which can flow laterally, with a consistent water table in both domains (Shen and Phanikumar, 2010). In this basin, with the exception of some dry lake beds, the groundwater table is so deep that most of the soil water flow is similar to a free-drainage boundary condition. The many years required for recharge to reach the water table in this basin is a major practical obstacle. Therefore, we recorded the flux that travels downward through the cell interface five meters bgs as the direct soil-column recharge R that eventually reaches the water table. While at local scales there may be (discontinuous) clay layers that impede vertical flow, we are concerned with large-scale, long-term-average fluxes.

The mountain-front subsurface recharge, Q_{Msub}, is added as a source term to the MODFLOW model for the valley aquifer. As discussed earlier, on the mountains, there is only a 30 cm soil layer. As water infiltrates into the mountain soil and accumulates above the bedrock, it can flow laterally in this layer in the model as governed by the 2D saturated groundwater flow equation. As the water flows from the mountainous region into the valley, it is treated as the mountain-front subsurface recharge. While lateral flow occurs, the water is at the same time drawn to the surface due to capillary pressure and evaporation.

In general, K estimated by MODFLOW+PEST calibration can be fed to PAWS+CLM in an iterative loop (the dashed line in Figure 4b). However, in our case, the groundwater is very deep in most regions, so this coupling is weak. Using different K fields in PAWS+CLM makes little difference in the valley floor.

279 *2.3.2 Configurations of the numerical models*

280 For domain discretization of the PAWS+CLM model, we use an 800 x 800 m^2 horizontal grid
281 size for the area shown in Figure 2a. Forty vertical layers, which are exponentially finer near the
282 surface, span the unsaturated zone between the ground surface and top of the confined aquifer.
283 We fitted a linear model to the sonic-porosity data to set porosity (θ_s) as a function of depth. As
284 described in Shen et al. (2014), we incorporated a 30 m resolution digital elevation model,
285 land-use data, and nationally maintained weather station data, along with our *in situ*
286 meteorological stations. Weather data from January 1, 2002, to December 31, 2016, were used to
287 drive the PAWS+CLM model. To spin-up the model, we first cycled it three times through the
288 forcing data from 2002 to 2016. Then, we extracted simulated recharge from 2006 to 2016 to be
289 used in MODFLOW, which was assumed to continue from the end of 2016.

290 Two layers of aquifers are represented in our model: the alluvium is the first (upper) layer and
291 the Bouse Formation, intermixed with Fanglomerate, is the second (underlying) layer. We pieced
292 together multiple sources of information, including borehole data, well logs, and gravity-based
293 surveys to obtain a consistent hydrostratigraphic framework for describing the valley subsurface.
294 Figure 5 shows the depth to the basement rock (i.e., the bottom of the second modeled aquifer
295 layer) resulting from this fusion of information as described below. For the second layer, a
296 buried ridge (drawn in Figure 5) is set as the rough western boundary of the Bouse Formation.
297 This ridge separates the upper and lower Chuckwalla valleys. In the lower valley, we assumed a
298 constant top elevation for the second layer (Bouse), since, as a marine/estuarine formation, the
299 Bouse is observed to have a flat surface. In the lower valley, the alluvium aquifer has a
300 maximum thickness of 165 m, and the Bouse 950 m. In the upper valley, there is no clear divide
301 between the alluvium and the lacustrine silt/clay layer, nor is there detailed data coverage.

Therefore, a constant thickness of ~90 m was assumed for the top (alluvium) layer (this value was obtained from geophysical surveys along a transect—GEI, 2010). With respect to the bottom of the second model layer, east of the divide, we constructed the bottom of the Bouse/Fanglomerate layer in the lower basin by combining gravity modeling (GEI, 2010) using Bouguer gravity data (Mariano et al., 1986) and well depths reaching the bedrock (Appendix 1 in WorleyParsons (2009)). West of the buried ridge, the bedrock map was reproduced from Figure 6 of Appendix C in GEI (2010), which is estimated by Bouguer gravity from GeoPentech. These data show that there are several pockets in the central part of the valley where the Bouse/Fanglomerate is very deep (>1000 m), forming bowl-like contours. However, because water in the deep part of the "bowls" does not participate in regional groundwater flow, we assumed a maximum depth of 1000 m.

Later in this paper, we report the recharge inside a "water mass-balance mask" (Figure 2b). This is different from the watershed boundary, in that some corner regions in the topographic basin are excluded from this mask, although all the simulated recharge in the PAWS+CLM model domain was extracted and provided to MODFLOW-PEST. The recharge in excluded areas has little impact on the groundwater level in the valley and thus should not be considered when reporting water budgets. Some of these regions could have a small groundwater exchange with adjacent valleys. We tracked the inflow from the Pinto Basin, which was the only basin suspected to have significant input. The agricultural region in the east and the Pinto Valley in the northwest are not included in the calibrated groundwater flow model, because fixed-head boundaries were used. The underflow through the southwest boundary should be minor and was not modeled.

For ordinary PAWS+CLM simulations, the channel network can be extracted from a stream network dataset such as the National Hydrography Dataset (Shen et al., 2014) or derived from

the flow network (Niu et al., 2017). Then the overland flow and groundwater contribute to the streamflow. However, in the study basin, the only river with a significant amount of flow time is the Colorado River at the eastern boundary, which has hydraulic connections to the local aquifers. All other flow paths are ephemeral washes. These washes are very small in width, and no data can comprehensively describe them. Thus, these flow paths were modeled as overland flow, which is naturally driven by topographic convergence.

2.3.3 Soil parameter adjustment

The van Genuchten (1980) water retention formula is written as $S = \frac{\theta(\psi)-\theta_r}{\theta_s-\theta_r} = (1 + |\alpha\psi|)^{-(N-1)/N}$, where S is relative saturation, ψ is the pressure head, θ is the moisture content, θ_r is the residual moisture content, and θ_s is the saturated moisture content (porosity). Calibrated soil parameters are α, N, θ_r, and θ_s. The unsaturated vertical conductivity at the relative saturation S is calculated by $K_z(S) = K_S S^{\lambda}\left[1 - \left(1 - S^{N/(N-1)}\right)^{(N-1)/N}\right]^2$, where the saturated conductivity K_S and the parameter λ are also calibrated. We manually adjusted these parameters on a trial and error basis, by matching the daily simulated soil moisture (with climate forcing from nationally maintained weather stations, as shown in Figure 2) to the *in situ* measurements at two SCAN sites between Nov. 2011 and Feb. 2016. The parameters were estimated separately for the Desert Center site, which represents the alluvial deposit soil, and the Ford Dry Lake, which represents the dry lake playa. The calibrated parameters are applied to the different zones in the PAWS+CLM simulation. We tried to match not only the moisture peaks but also interpeak minima. After suitable adjustment factors (multipliers and additions) were found, we applied the parameters to their respective soil zones. Across different layers, all parameters except residual moisture content were kept uniform.

2.4. Calibration of groundwater conductivity using MODFLOW+PEST

Although PAWS+CLM already contains a groundwater model, we utilized the MODFLOW+PEST package (Doherty, 2003; Tonkin and Doherty, 2009) to calibrate the *K* field, because it provides greater flexibility for representing a heterogeneous *K* field. Essentially, this package allows users to insert scattered "pilot points", whose conductivity values are the parameters to be adjusted by the PEST algorithm. These pilot points provide the handles for enforcing parameter changes as well as spatial relationships. The conductivity field in the rest of the domain is interpolated from the calibrated values at the pilot points based on the spatial relationship. In terms of the spatial relationships, both kriging and inverse distance weighting (IDW) with different options (e.g., ordinary or simple kriging, different methods of IDW) can be specified as models representing geostatistical constraints. These constraints are implemented in the form of regularization terms, i.e., a penalty term to the *K* values so these values do not deviate significantly from what is inferred from the geostatistical model. For kriging, the empirical variogram needs to be specified *a priori*. In contrast, the parameters in IDW can be estimated along with the *K* field. Since we have a very limited set of known *K* estimates which prevented us from building an *a-priori* variogram, we employed IDW and allowed estimation of the appropriate IDW parameters. Because of the limited number of parameters, IDW-based calibration always converged.

Normally, for an integrated hydrologic model like PAWS+CLM, the parameter sets are only adjusted by applying a multiplier; that is, the spatial heterogeneity is preserved and all conductivity values are multiplied by or added to a constant. The newly proposed framework allows the calibration of a spatially distributed *K* field with imposed geostatistical constraints. It should be possible, with much additional work, to implement PEST directly within PAWS and

completely avoid MODFLOW. However, here we used a "soft" coupling approach that is more generic, in that either PAWS or PEST may be replaced by the modeler's choosing. The present approach is also simpler and can be carried out by practitioners and researchers. PAWS and CLM are integrated at the source code level, while recharge from PAWS+CLM was extracted from its outputs and provided as input files to the MODFLOW model (Figure 4). A 2-layer MODFLOW model of the Chuckwalla Basin was set up for the valley portion of the basin. The MODFLOW model had a horizontal grid spacing identical to PAWS+CLM. MODFLOW+PEST was used to calibrate the *K* fields to steady-state water-table levels in observation wells. Constraining the possible range of *K* is important for reducing overfitting, which means *K* is adjusted unrealistically to fit the noise rather than the true signal. For the top aquifer layer, we added pumping-test-estimated *K* as known values and constrained *K* between 0.1 to 30 m/day. For the second layer, understanding that pumping tests are rarer and most *K* estimates are within 1.5 to 4 m/day, we constrained the conductivity in the range of 0.1 to 6 m/day.

2.4.1 Groundwater withdrawals and boundary conditions

Presently, a prison and a resort pump about 7100 m^3/day (2.6 Mm^3yr^{-1}) and 3684 m^3/day (1.3 Mm^3yr^{-1}) from the Bouse and the alluvium formation, respectively (WorleyParsons, 2010) (Figure 2). These sink terms have existed for over two decades, and they have been included for calibrating the steady-state model. For future projections, we added recently approved water users and the proposed EMPS Project, as shown in Table 1, with water-use data taken from their respective project EIA reports.

The eastern boundary of the MODFLOW model ends at the western perimeter of the Palo Verde Mesa agricultural zone (Palo Verde Valley), where USGS well data are available to build a fixed-head boundary condition to avoid modeling irrigation and withdrawals (Figure 2b).

Mountain boundaries of the MODFLOW model are set as no-flow boundary conditions, but as discussed earlier, mountain-front subsurface flow into the valley is added as recharge. The Pinto Basin connects to the Chuckwalla Basin through a thin sedimentary neck (Figure 2b). No groundwater observations in the Pinto Valley were readily available, so we used an average *K* value there in PAWS+CLM and excluded it from calibration to reduce the number of parameters and overfitting. PAWS-Simulated inflow from the Pinto Basin, which varied across different perturbed simulations, was added as a source term to the Chuckwalla Basin. The Orocopia Valley Basin parallels the Interstate 10 transportation corridor to the southwest. The underflow to the Orocopia Valley Basin was deemed negligible.

2.5. Ensemble simulations, model rejection, and the dual-model integration

Our goal of assembling an ensemble of simulations was not to estimate the probability distribution of withdrawal impacts, but to put bounds on such impacts given large parametric uncertainties. We first identified several key uncertain soil parameters (Table 3) for which preliminary experiments showed strong impacts on recharge. We also tested a parameter describing vegetation interception of runoff, but it was not found to be a sensitive parameter, likely because most of the recharge runs off from mountains where there is little surface cover. Then we perturbed the parameters simultaneously, using global multipliers to generate recharge maps ranging from high to low (Figure 1). Higher recharge values lower the impact of pumping. The calibrated soil parameters served as the base case (#6) in these simulations. For each recharge estimate, we calibrated the model five times by assuming different (but spatially constant) initial *K* values. Initial values from 3 to 7 m/day were used to seed the calibrations.

Four tests were used to reject *K* fields: (1) a field is rejected if the calibrated head has significant bias from observations despite the calibration, assessed using a z-test of the mean of the residuals

(simulated minus observed groundwater head). The residuals should have a zero mean, i.e., a bias of nearly zero, when the amount of recharge is in an approximately admissible range. This is typically the easiest objective to achieve during the calibration, for a bias will always produce a large squared loss. A large bias could only mean the algorithm has truly failed to find an appropriate *K* field that could lower the error, even with a large number of degrees of freedom. A z-test examines if the bias is large enough so that the residuals could not have come from a zero-mean Gaussian distribution with a specified variance. For this variance, to be lenient (i.e., to use more relaxed rules for model rejection), we used four times the variance from the best-calibrated case; (2) we conducted a Chi-squared test on the residual variance, which detects if the variance is too large, also using four times the variance from the best-calibrated case. When the residual variance was too large, the calibrated head could show too much or too little variation compared to the observations, typically resulting from an incorrect spatial distribution of recharge; (3) we ran a regression test between elevation and residual. If residuals are significantly correlated to elevation, there is a regional pattern to errors, implying errors are spatially auto-correlated, and the residual independence assumption is violated. This correlation results from regionally biased recharge, which leads to overfitting during calibration; (4) when the PEST calibration overfits to observations, it tends to introduce unrealistically large spatial variation in the *K* field in a small neighborhood. To detect this issue, a bi-quadratic surface was fitted to the *K* field, in order to calculate the standard deviation for the *K* residual from the surface. Cases with large standard deviation in *K* were rejected. However, as we shall see below in Results and Discussion, no recharge field was rejected due to this criterion. Five calibrations were conducted for each recharge case shown in Table 3, using different initial guesses of *K*.

To project future transient groundwater table changes arising from pumping, we ran transient MODFLOW models using extracted recharge maps from different scenarios as a constant input, while transient pumping was applied according to the schedules in Table 1, assuming 2017 was the starting year. The spatial resolution of these projection runs was kept the same as the calibration runs. These projections assumed that the long-term-averaged recharge would not change substantially in the future, even though pumping would. Another issue with making future projections of pumping impact is that we also needed the aquifer storage properties, but steady-state calibrations do not constrain storage parameters. We thus ran an ensemble of transient 24-year-long simulations, continuing from 2017, using different sets of storage parameters to examine their impact on groundwater sustainability. We considered the plausible ranges of the specific yield of the alluvium (S_y) and the specific storage of the lower layer (S_s). For S_y we tested {0.05, 0.10, 0.15} (dimensionless). A small value of 0.05 was estimated for Desert Center (WorleyParsons, 2009). However, other estimates place the value around 0.15. For S_s, earlier studies of aquifers in this area have bounded the range from 5×10^{-6} to 1×10^{-4} (m^{-1}), so three values were tested in this study: {1×10^{-6}, 5×10^{-6}, 5×10^{-5}} (m^{-1}).

3. Results

3.1. Soil moisture comparisons

After the soil parameters (Table 3) were adjusted, the Richards Equation-based PAWS+CLM model was able to match the soil-moisture time series at both stations (Figure 6). The Correlation Coefficients reported in Figure 6 were higher than those obtained in a Midwest basin (Riley and Shen, 2014). At the Desert Center site, the simulated soil moisture matched very well with the observations at 5 cm and 10 cm depths. At 5 cm, there was a slight underestimation of peaks, while the minimum soil moisture was slightly higher. At this site, the 20 cm probe appeared to

have malfunctioned: it recorded moisture rises that were much larger than those detected at the surface. Therefore, it is not shown here. At 50 cm depth, while the timing of the moisture wave was not completely correct, the amplitude of seasonal fluctuation was similar between observed and simulated. The interannual climb of soil moisture from 2012 and 2013 was correctly described by the model, showing that a slow infiltration over the period of a year could be captured. At Ford Dry Lake, peaks were similarly well captured. The 20 cm probe also captured the peaks, but the simulated moisture lagged behind the observed in the 2014 rain event. The calibrated K_S values are around 0.1 m/day at both sites (Table 2), which is lower than the expected range for sandy soils. This value is in the low range of the values reported for Mojave Desert soils, which were measured between 0.07 to 350 m/day for old and young soils, respectively (Young et al., 2004). However, despite some large rainfall events, the observed moisture content seldom rose above 0.15, and spent the majority of the time below 0.05 (Figure 6). Therefore, the unsaturated conductivity, which can be orders-of-magnitude lower than K_S, played a more important role in limiting infiltration than K_S. Hence, the van Genuchten parameters were more influential than K_S for estimating infiltration and recharge. Their adjustment could have compensated for uncertainties in K_S.

3.2. Assessing and rejecting perturbed simulations

Five of the recharge fields, which were near either the high or low end of recharge rates from the simulations, were rejected as a result of their inability to fit the groundwater head (Table 4 and Figure 7). The most frequent criterion that caused rejection of the cases was the bias (a z-test between the mean of the observed and calibrated groundwater heads). This test rejected all the cases for recharge #1,-#5, and #12. Moreover, every calibration case from recharge #1-#5 and #12 was also rejected by more than one condition. Recharge estimates from simulations #1

through #5 (simply called recharge #1-#5—similarly below) were rejected by all tests as they overestimated the groundwater head (Figure 8), suggesting that their values were too large. In contrast, recharge #12 underestimated the groundwater head regardless of calibration, suggesting that its value was too low. As described in Section 2.5, we implemented relaxed rejection criteria for three statistical tests, using a confidence level of 2% and an assumed variance that is four times that of the best calibrated field (var_{min}, from recharge #10, realization 2). Simulation #11 case 5 was a borderline case. It was the only retained case from simulation #11, and it would have been rejected if we had used 2.25 times var_{min} as our rejection criteria. Therefore, we labelled #11 as "unlikely." We tried increasing the soil conductivity on the mountains by a factor of 2 in simulation #7, but these simulations were also rejected. Figure 8 presents the observed and calibrated groundwater head for some examples of accepted and rejected simulations. The elevation-regression test rejected some cases for recharge #6 through #10. The K-field variance test by itself did not reject any cases.

Using recharge generated by the default parameter set (simulation #6), we found that the spatially distributed hydraulic head compared well with the observations (Figure 8), with only a few meters of difference at the maximum for each data point. However, for recharge #4, the groundwater head was always overestimated, regardless of the calibration effort and the initial guess for K. Recharge #5 overestimated head in the lower basin (where observed head is about 80 m in Figure 8). Simulation #11 case 2, on the other hand, apparently underestimated groundwater head, suggesting too little recharge. Overall, the magnitude and variation of *K* conformed to our knowledge of the area.

With recharge #6, the simulated groundwater contours (Figure 9a) were in agreement with trends shown in earlier EIA (WorleyParsons, 2010). The water table followed the topographic trend and descended from the northwest to the southeast, showing a smaller gradient than the topographic slope. The groundwater table had a larger gradient in the upper Chuckwalla valley than in the lower, due to the inflow from the Pinto Basin. The gradient was small in the wide valley floors due to little recharge in the valley floor, but larger in the corridor that connects the upper and lower valleys. Figure 2a shows only two groundwater wells in the corridor. As a result, this part of the model domain was not well constrained. This pattern resulted from small calibrated *K* values at the corridor and larger *K* in the main valleys, especially toward the eastern boundary. The calibrated *K* field was mostly smooth (Figure 9b) except for a small high-conductivity zone in the north-central corner.

3.3. Water balance of the basin under uncertainty

Since most of the calibrations with recharge #11 were rejected, and only one was narrowly retained, we took the average recharge value of #10 and #11 as the lower bound estimate of total inflow (recharge + cross-valley underflow), which was 3.07 mm/yr (8.8 Mm^3yr^{-1}). The upper bound of our inflow estimate was 4.99 mm/yr (14.2 Mm^3yr^{-1}), taken from recharge #6. Our estimates ranged from 3.4% to 5.6% of precipitation. Our lower-bound value is close to the high-range extrapolated estimates from the National Park Service, while our upper-bound estimate is similar to the Maxey-Eakin-based estimates from previous EIA (FERC, 2012; Godfrey et al., 2012). In the literature, recharge estimates in arid and semi-arid basins in the southern Mojave ranged from 3%-7% of precipitation (Stonestrom et al., 2007). Reported values in nearby basins ranged from 2.8%-5.2% (Whitt and Jonker in CGB 2004), down to 1.1% (Nishikawa et al., 2005).

Simulated recharge from the default case was focused in ephemeral washes and alluvial fans on mountain fringes (Figure 10a). Note that the largest recharge occurred along washes in annotated regions A and B, which agrees with the vegetation pattern seen from satellite images in Figure 3c. At the mountain edges, we note mostly positive recharge due to run-on infiltration of overland flow running off the mountains, then infiltrating through the soils of the alluvial fan (either via flow domain leakance or infiltration of backfilled water through the soil matrix — backfill occurs when water level in the conceptualized flow compartment is higher than that in the ponding compartment). After large storms, as runoff passes through the alluvial fans, water could leak via the overland flow channel as Q_{og} or flood the ponding domain and infiltrate as soil-column recharge (Section 2.3.1). The thick sediment underneath the alluvial fans provided more volume for storage and infiltration. However, there were also some negative values that resulted from groundwater return flow from the thin mountain soils. There was also negative recharge near Ford Dry Lake (near the location of the FDL-site in Figure 2a), where groundwater emerges and evaporates away. Q_{Msub}, the lateral groundwater flow via the thin soils on the mountains, took place exclusively at the feet of mountains (Figure 10b). The magnitude of Q_{Msub} appeared much smaller than the total recharge. This flux turned out to play a minor role due to the limited thickness of the simulated mountain soils.

3.4. Projections of the impacts of pumping on groundwater sustainability

Recharge estimates from the retained simulations and their respective calibrated *K* fields were used to estimate drawdown in response to the pumping due to new solar-plants. At the EMPS site, the largest drawdown occurred at the end of the initial fill period and had a range of 8 to 11 meters when $S_y = 0.05$ (Figures 11a and 12). Without rejection of the overfitted simulated recharge rates, this range would have been 7 to 15.3 meters. The reduction of uncertainty

depended on the location. At some water user sites (EMPS and Desert Sunlight, Figures 11a and 11b), the possible range of drawdown has been significantly narrowed. However, at another site (Genesis, Figures 11c), the model rejection did not greatly reduce the uncertainty, since there were fewer groundwater observations to constrain the model around this site.

For EMPS at $S_y = 0.05$, the water table declined by around 3 m within one year after the initial-fill phase. Heavy pumping induced a large hydraulic gradient and a deep cone of depression. Once initial-fill pumping ceased, the large aquifer transmissivity allowed groundwater flow to rapidly fill the cone. The water table then gradually declined during the project's refill phase. When the pumping at the EMPS site was stopped, the water table recovered 4 m in one year, and after 20 years of simulation, the water table recovered to 6-7 m below initial values. According to this trajectory, the system may appear to recover quickly from the assumed pumping, but the large initial recovery speed does not imply that it can return to pristine conditions. If there is a boom of new water users, groundwater levels will not be sustainable.

Figure 12 illustrates the effect of S_y on groundwater drawdown as a result of pumping. Aquifer drawdown was more sensitive to the assumption of S_y than the recharge employed (Figures 12), which highlights the importance of obtaining more accurate estimates of S_y. However, the drawdown was not sensitive to specific storage (S_s) in the range tested. Hence, the results from the S_s sensitivity tests are not shown here. To the west of Desert Center, the first aquifer layer (with a thickness of approximately 13 m) dried after pumping. Recharge #6 was the highest accepted recharge rate, while recharge #10, which resulted in a deeper cone of depression, was the lowest accepted recharge.

If, as in conventional methods, we had assumed a uniformly distributed recharge before calibrating *K*, errors would be introduced into different parts of the basin, even with the same total recharge. The uniform recharge tended to overestimate the groundwater head in the lower basin even for recharge #10 (Figure 13a-b). While the root-mean-square error was not very high, the resulting *K* fields had a higher local variation. Using a uniform recharge had spatially heterogeneous impacts on drawdown: the range of drawdown for retained simulations at EMPS was larger than using a spatially distributed recharge, while the opposite was true at the Genesis site (Figure 13c-f). In addition, at the Genesis site, using a uniform recharge led to an underestimation of pumping drawdown. This difference results from the EMPS Project being closer to the mountain front and wash recharge, while Genesis is located in the valley center, far from the focus of recharge. Thus, uniform recharge overestimated recharge near Genesis.

A previous report utilized uniformly distributed recharge and PEST to estimate *K* and the effects of pumping (Greer et al., 2013). Their calibration resulted in a very large variation in the *K* field and 30 meters of drawdown (Figure 9 in their paper). We believe such drawdowns were overestimated, resulting from uniformly imposed recharge.

4. Discussion

In the past, it has been difficult to simultaneously incorporate both soil moisture and spatially distributed groundwater data within modeling. Recharge estimates cannot be directly constrained by groundwater head data, because the spatial distribution of *K* can significantly influence the simulated groundwater head, and we do not have an extensive prior estimate of *K*. Allowing the spatial calibration of *K* using PEST increases the model's ability to adapt to recharge and compensate for some errors in recharge estimations. Nonetheless, we can see that if recharge rates were too high or too low, the resulting simulation would not be admissible even if

596 calibration were allowed. Our method is therefore a more lenient way of reducing uncertainty,
597 which we adopted considering the challenges of modeling the desert, the scarcity of available
598 information, and associated uncertainties. Our proposed dual-model approach, which should be
599 applicable to other arid regions proves to be effective in identifying a plausible range of recharge
600 values for mountainous desert regions. If within a region there are recharge terms that are
601 omitted or overestimated, e.g., due to local clay impedance or leakance from a water conduit,
602 simulations with perturbed parameters can (to some extent) compensate for the error. Eventually,
603 only a recharge rate within a suitable range can pass the validation using groundwater
604 observations. The calibrated *K* field significantly influences possible drawdown and recovery,
605 which is also why the integration of groundwater observations is useful.

606 Previous research on recharge in arid regions has heavily focused on infiltration beneath washes.
607 Our study suggests that an overlooked area for potential recharge is the alluvial fans. As
608 immediate recipients of mountain runoff, the fans and adjacent flat areas have the first chance to
609 hold and infiltrate water. While some chloride studies suggested little deep recharge under some
610 fans (Stonestrom et al., 2004), other field studies (Bull, 1977; Houston, 2002) and modeling
611 studies (Blainey and Pelletier, 2008; Munévar and Mariño, 1999) found alluvial fans to be highly
612 local but significant recharge areas. Our results suggest relevant data, e.g., moisture or solute
613 concentration under alluvial fans, are needed to better verify modeling results.

614 Water managers may find fast water table recovery to be reassuring and sometimes use it as a
615 guideline to manage water. However, heavy pumping induces a steep cone of depression and
616 thus a large hydraulic gradient that would be automatically followed by a rapid initial recovery
617 after pumping cessation. Therefore, the speed of recovery itself cannot guarantee sustainability,
618 because the water may not ultimately recover to pre-pumping levels.

5. Conclusion

We have proposed a novel dual-model approach to provide a bounded estimate of the effects of groundwater pumping in a mountainous arid region. An integrated surface-subsurface hydrologic model can better approximate the locations and distribution of recharge, while incorporation of groundwater head data is crucial for constraining the recharge rates. Our results indicate that conventional approaches assuming uniform recharge will distort the calibrated *K* field and yield very different projections. Given the estimated range of recharge, groundwater levels will likely decrease across the Chuckwalla Basin over the life of new large-scale water users. Once pumping ceases, groundwater levels may initially recover quickly, but ultimately not to pre-pumping levels. Our results also highlight the importance of obtaining accurate estimates of aquifer storage parameters when evaluating the impacts of pumping.

6. Limitations

This work should not be a replacement for detailed field measurements of recharge, such as chemical and isotopic analysis of samples. Owing to the extensive amount of work needed to cover sporadic recharge, completion of fieldwork can be challenging. The present work provides one way of estimating recharge in this setting, which is constrained by a model, as well as groundwater observations, and should be considered together with other estimates.

Compared to the DA approach, the present approach does not provide a formal estimate of uncertainties. DA-based approaches that enforce geostatistical constraints such as MPS and variograms, which have only recently been proposed, could be attempted in the future. However, in arid mountainous regions with scattered data sources, there might be fresh new challenges, such as nonconvergence issues.

Acknowledgments

We appreciate Kristan Culbert and James Collins for compiling well data. This work was supported by the U.S. Bureau of Land Management under Interagency Agreement L11PG00354, as part of Work for Others funding from Lawrence Berkeley National Lab, provided by the U.S. Department of Energy, Office of Science, under Award Number DE-AC02-05CH11231. This paper does not represent the position of the United States government. The present work was partially developed within the framework of the Panta Rhei Research Initiative.

References

Blainey, J.B., Pelletier, J.D., 2008. Infiltration on alluvial fans in arid environments: Influence of fan morphology. J. Geophys. Res. 113, F03008. https://doi.org/10.1029/2007JF000792

BLM, 2017. Palen Solar Photovoltaic Project: Informational update related to the future draft supplemental EIR/EIS, Alternatives Overview [WWW Document]. Bur. L. Manag. , Palm Springs, Palm Springs-South Coast F. Off. Calif. URL https://eplanning.blm.gov/epl-front-office/projects/nepa/68122/99192/120213/PalenOverview_with_signature(508)_FINAL.pdf (accessed 10.8.17).

BLM, 2011. Desert Sunlight Solar Farm Project California Desert Conservation Area Plan Amendment and Final Environmental Impact Statement [WWW Document]. Bur. L. Manag. CACA #48649. URL https://energy.gov/nepa/downloads/eis-0448-final-environmental-impact-statement (accessed 12.31.14).

Bull, W.B., 1977. The alluvial-fan environment. Prog. Phys. Geogr. 1, 222–270. https://doi.org/10.1177/030913337700100202

CADWR, 1979. Sources of Powerplant Cooling Water in the Desert Area of Southern California -- Reconnaissance Study [WWW Document]. Calif. Dep. Water Resour. Bull. 91-24. URL http://www.water.ca.gov/waterdatalibrary/docs/historic/Bulletins/Bulletin_91/Bulletin_91-24__1979.pdf (accessed 7.29.16).

CGB, 2004. Copper Mountain Valley Groundwater Basin [WWW Document]. Calif. Groundw. Bull. 118. URL http://www.water.ca.gov/groundwater/bulletin118/basindescriptions/7-11.pdf (accessed 1.1.17).

Doherty, J., 2003. Ground water model calibration using pilot points and regularization. Groundwater 41, 170–7.

EPA, 2012. Draft Environmental Impact Statement for the Proposed Desert Harvest Solar Project, Riverside County, California (CEQ #20120099) [WWW Document]. U.S. Environ. Prot. Agency (EPA), Detail. comments Draft Environ. impact statement Propos. Desert Harvest Sol. Proj. Riverside County, California, July 13, 2012. URL https://archive.epa.gov/region9/nepa/web/pdf/desert-harvest-solar-deis.pdf

Erdal, D., Cirpka, O.A., 2016. Joint inference of groundwater-recharge and hydraulic-conductivity fields from head data using the ensemble Kalman filter 555–569. https://doi.org/10.5194/hess-20-555-2016

Everett, R.R., 2013. Chuckwalla Valley Multiple-well Monitoring Site, Chuckwalla Valley, Riverside County, California [WWW Document]. US Geol. Surv. Open-File Rep. 2013–1221. URL https://pubs.usgs.gov/of/2013/1221/ (accessed 5.24.17).

Fan, Y., Li, H., Miguez-Macho, G., 2013. Global Patterns of Groundwater Table Depth. Science (80-.). 339.

FERC, 2012. Final Environmental Impact Statement for the Proposed Eagle Mountain Pumped Storage Hydroelectric Project (P-13123-002) [WWW Document]. Fed. Energy Regul.

690 Comm. URL https://www.ferc.gov/industries/hydropower/enviro/eis/2012/01-30-12.asp
691 (accessed 5.23.17).

692 Flint, A.L., Flint, L.E., Hevesi, J.A., Blainey, J.B., 2004. Fundamental concepts of recharge in the
693 desert southwest: A regional modeling perspective, in: Hogan, J.F., Phillips, F.M., Scanlon,
694 B.R. (Eds.), Groundwater Recharge in a Desert Environment: The Southwestern United
695 States. American Geophysical Union, Washington, DC, pp. 159–184.
696 https://doi.org/10.1029/009WSA10

697 GEI, 2010. Environmental Impact Analysis: Groundwater [WWW Document]. Eagle Mt. Pumped
698 Storage Proj. Draft Environ. Impact Rep. Vol. I. URL
699 http://www.waterboards.ca.gov/waterrights/water_issues/programs/water_quality_cert/eagle
700 mtn13123_eir.shtml (accessed 8.1.16).

701 Godfrey, P., Ludwig, N., Salve, R., 2012. Groundwater and large-scale renewable energy projects
702 on federal land: Chuckwalla valley groundwater basin, in: Arizona Hydrological Society
703 2012 Annual Water Symposium.

704 Greer, C., Quinn, J.J., Carr, A., O'Conner, B., 2013. A groundwater model to assess water resource
705 impacts at the Riverside East Solar Energy Zone [WWW Document]. Rep. Prep. Bur. L.
706 Manag. URL
707 http://blmsolar.anl.gov/sez/ca/riverside-east/groundwater/downloads/Riverside-East-Ground
708 water-Report.pdf

709 Gunduz, O., Aral, M.M., 2005. River networks and groundwater flow: a simultaneous solution of a
710 coupled system. J. Hydrol. 301, 216–234. https://doi.org/10.1016/j.jhydrol.2004.06.034

711 Harbaugh, A.W., 2005. MODFLOW-2005, the U.S. Geological Survey modular ground-water
712 model -- the Ground-Water Flow Process: U.S. Geological Survey Techniques and Methods
713 6-A16. Reston, Virginia.

714 Houston, J., 2002. Groundwater recharge through an alluvial fan in the Atacama Desert, northern
715 Chile: mechanisms, magnitudes and causes. Hydrol. Process. 16, 3019–3035.
716 https://doi.org/10.1002/hyp.1086

717 Izbicki, J.., Radyk, J., Michel, R.., 2000. Water movement through a thick unsaturated zone
718 underlying an intermittent stream in the western Mojave Desert, southern California, USA. J.
719 Hydrol. 238, 194–217. https://doi.org/10.1016/S0022-1694(00)00331-0

720 Ji, X., Shen, C., 2018. The introspective may achieve more: enhancing existing Geoscientific
721 models with native-language structural reflection. Comput. Geosci. 110.
722 https://doi.org/10.1016/j.cageo.2017.09.014

723 Ji, X., Shen, C., Riley, W.J., 2015. Temporal evolution of soil moisture statistical fractal and
724 controls by soil texture and regional groundwater flow. Adv. Water Resour. 86, 155–169.
725 https://doi.org/10.1016/j.advwatres.2015.09.027

726 Knowling, M.J., Werner, A.D., 2016. Estimability of recharge through groundwater model
727 calibration : Insights from a field-scale steady-state example. J. Hydrol. 540, 973–987.
728 https://doi.org/10.1016/j.jhydrol.2016.07.003

729 Mariano, J., Helferty, M.G., Gage, T., 1986. Bouger and Isostatic Residual Gravity Maps of the

730 Colorado River Region, including the Kingman, Needles, Salton Sea, and El Centro
731 Quadrangles [WWW Document]. Open-File Rep. 86-347. URL
732 https://pubs.er.usgs.gov/publication/ofr86347 (accessed 6.12.16).

733 Markstrom, S.L., Niswonger, R.G., Regan, R.S., Prudic, D.E., Barlow, P.M., 2008.
734 GSFLOW—Coupled Ground-Water and Surface-Water Flow Model Based on the
735 Integration of the Precipitation-Runoff Modeling System (PRMS) and the Modular
736 Ground-Water Flow Model (MODFLOW-2005) [WWW Document]. U.S. Geol. Surv.
737 Chapter 1 Sect. D, Ground-Water/Surface-Water B. 6, Model. Tech. URL
738 http://pubs.usgs.gov/tm/tm6d1/

739 Maurer, D.K., Berger, D.L., 2006. Water Budgets and Potential Effects of Land- and Water-Use
740 Changes for Carson Valley, Douglas County, Nevada, and Alpine County, California [WWW
741 Document]. USGS Sci. Investig. Rep. 2006–5305. URL
742 https://pubs.usgs.gov/sir/2006/5305/pdf/sir20065305.pdf (accessed 1.1.17).

743 Maxey, G.B., Eakin, T.., 1949. Groundwater in the White River Valley, White Pine, Nye, and
744 Lincoln counties, Nevada [WWW Document]. Water Resour. Bull. No. 8, State Nevada, Off.
745 State Eng. URL https://www.nrc.gov/docs/ML0331/ML033140348.pdf

746 Maxwell, R.M., Condon, L.E., Kollet, S.J., 2015. A high-resolution simulation of groundwater and
747 surface water over most of the continental US with the integrated hydrologic model ParFlow
748 v3. Geosci. Model Dev. 8, 923–937. https://doi.org/10.5194/gmd-8-923-2015

749 Maxwell, R.M., Putti, M., Meyerhoff, S., Delfs, J.-O., Ferguson, I.M., Ivanov, V., Kim, J., Kolditz,
750 O., Kollet, S.J., Kumar, M., Lopez, S., Niu, J., Paniconi, C., Park, Y.-J., Phanikumar, M.S.,
751 Shen, C., Sudicky, E.A., Sulis, M., 2014. Surface-subsurface model intercomparison: A first
752 set of benchmark results to diagnose integrated hydrology and feedbacks. Water Resour. Res.
753 50, 1531–1549. https://doi.org/10.1002/2013WR013725

754 McFadden, L.D., Wells, S.G., Jercinovich, M.J., 1987. Influences of eolian and pedogenic
755 processes on the origin and evolution of desert pavements. Geology 15, 504.
756 https://doi.org/10.1130/0091-7613(1987)15<504:IOEAPP>2.0.CO;2

757 Metzger, D.G., Loeltz, O.J., Irelna, B., 1973. Geohydrology of the Parker-Blythe-Cibola Area,
758 Arizona and California, United States Geological Survey Professional Paper 486-G.
759 Washington, D. C.

760 Mirus, B.B., 2015. Evaluating the importance of characterizing soil structure and horizons in
761 parameterizing a hydrologic process model. Hydrol. Process. 29, 4611–4623.
762 https://doi.org/10.1002/hyp.10592

763 Mirus, B.B., Perkins, K.S., Nimmo, J.R., Singha, K., 2009. Hydrologic Characterization of Desert
764 Soils with Varying Degrees of Pedogenesis: 2. Inverse Modeling for Effective Properties.
765 Vadose Zo. J. 8, 496. https://doi.org/10.2136/vzj2008.0051

766 Munévar, A., Mariño, M.A., 1999. Modeling Analysis of Ground Water Recharge Potential on
767 Alluvial Fans Using Limited Data. Ground Water 37, 649–659.
768 https://doi.org/10.1111/j.1745-6584.1999.tb01156.x

769 Nishikawa, T., Izbicki, J.A., Hevesi, J.A., Martin, C.L.S.P., 2005. Evaluation of Geohydrologic

Framework, Recharge Estimates and Ground-Water Flow of the Joshua Tree Area, San Bernardino County, California [WWW Document]. USGS Sci. Investig. Rep. 2004-5267. URL https://pubs.usgs.gov/sir/2004/5267/ (accessed 1.1.17).

Niu, J., Phanikumar, M.S., 2015. Modeling watershed-scale solute transport using an integrated, process-based hydrologic model with applications to cacterial fate and transport. J. Hydrol. 529, 35–48. https://doi.org/10.1016/j.jhydrol.2015.07.013

Niu, J., Shen, C., Chambers, J., Melack, J.M., Riley, W.J., 2017. Interannual variation in hydrologic budgets in an Amazonian watershed with a coupled subsurface - land surface process model. J. Hydrometeorol. https://doi.org/10.1175/JHM-D-17-0108.1

Niu, J., Shen, C., Li, S.-G., Phanikumar, M.S., 2014. Quantifying storage changes in regional Great Lakes watersheds using a coupled subsurface-land surface process model and GRACE, MODIS products. Water Resour. Res. 50, 7359–7377. https://doi.org/10.1002/2014WR015589

Oleson, K., Lawrence, D.M., Bonan, G.B., Flanner, M., Kluzek, E., Lawrence, P., Levis, S., Swenson, S., Thornton, P., Dai, A., Decker, M., Dickinson, R., Feddema, J., Heald, C., Hoffman, F., Lamarque, J., Mahowald, N., Niu, G., Qian, T., Randerson, J., Running, S., Sakaguchi, K., Slater, A., Stockli, R., Wang, A., Yang, Z., Zeng, X., 2010. Technical description of version 4.0 of the Community Land Model (CLM). NCAR Technical Note, NCAR/TN-478+STR., Boulder, Colorado.

Oleson, K.W., Lawrence, D.M., Bonan, G.B., Drewniak, B., Huang, M., Koven, C.D., Levis, S., Li, F., Riley, W.J., Subin, Z.M., Swenson, S.C., Thornton, P.E., Bozbiyik, A., Fisher, R., Heald, C.L., Kluzek, E., Lamarque, J.-F., Lawrence, P.J., Leung, L.R., Lipscomb, W., Muszala, S., Ricciuto, D.M., Sacks, W., Sun, Y., Tang, J., Yang, Z.-L., 2013. Technical Description of version 4.5 of the Community Land Model (CLM) [WWW Document]. NCAR/TN-503+STR, NCAR Tech. Note. URL http://www.cesm.ucar.edu/models/cesm1.2/clm/CLM45_Tech_Note.pdf

Owen-Joyce, S.J., Wilson, R.P., Carpenter, M.C., Fink, J.B., 2000. Method to identify wells that yield water that will be replaced by water from the Colorado River downstream from Laguna Dam in Arizona and California [WWW Document]. Water-Resources Investig. Rep. 00-4085. URL https://pubs.usgs.gov/wri/2000/4085/ (accessed 7.29.16).

Pasetto, D., Niu, G.-Y., Pangle, L., Paniconi, C., Putti, M., Troch, P.A., 2015. Impact of sensor failure on the observability of flow dynamics at the Biosphere 2 LEO hillslopes. Adv. Water Resour. 86, 327–339. https://doi.org/10.1016/J.ADVWATRES.2015.04.014

Pau, G.S.H., Shen, C., Riley, W.J., Liu, Y., 2016. Accurate and efficient prediction of fine-resolution hydrologic and carbon dynamic simulations from coarse-resolution models. Water Resour. Res. 52, 791–812. https://doi.org/10.1002/2015WR017782

Rasmussen, J., Madsen, H., Jensen, K.H., Refsgaard, J.C., 2016. Data assimilation in integrated hydrological modelling in the presence of observation bias. Hydrol. Earth Syst. Sci. 20, 2103–2118. https://doi.org/10.5194/hess-20-2103-2016

Rasmussen, J., Madsen, H., Jensen, K.H., Refsgaard, J.C., 2015. Data assimilation in integrated hydrological modeling using ensemble Kalman filtering: evaluating the effect of ensemble

size and localization on filter performance. Hydrol. Earth Syst. Sci. 19, 2999–3013. https://doi.org/10.5194/hess-19-2999-2015

Riley, W.J., Shen, C., 2014. Characterizing coarse-resolution watershed soil moisture heterogeneity using fine-scale simulations. Hydrol. Earth Syst. Sci. 18, 2463–2483. https://doi.org/10.5194/hess-18-2463-2014

RMC, 1990. Roscoe Moss Company (RMC) Handbook of ground water development. Wiley.

Scanlon, B.R., 2004. Evaluation of methods of estimating recharge in semiarid and arid regions in the southwestern U.S., in: Hogan, J.F., Phillips, F.M., Scanlon, B.R. (Eds.), Groundwater Recharge in a Desert Environment: The Southwestern United States. American Geophysical Union, Washington, DC., pp. 235–254. https://doi.org/10.1029/009WSA13

Schreiner- McGraw, A.P., Vivoni, E.R., 2018. On the Sensitivity of Hillslope Runoff and Channel Transmission Losses in Arid Piedmont Slopes. Water Resour. Res. 54, 4498–4518. https://doi.org/10.1029/2018WR022842

Shen, C., Niu, J., Fang, K., 2014. Quantifying the effects of data integration algorithms on the outcomes of a subsurface–land surface processes model. Environ. Model. Softw. 59, 146–161. https://doi.org/10.1016/j.envsoft.2014.05.006

Shen, C., Niu, J., Phanikumar, M.S., 2013. Evaluating controls on coupled hydrologic and vegetation dynamics in a humid continental climate watershed using a subsurface - land surface processes model. Water Resour. Res. 49, 2552–2572. https://doi.org/10.1002/wrcr.20189

Shen, C., Phanikumar, M.S., 2010. A process-based, distributed hydrologic model based on a large-scale method for surface–subsurface coupling. Adv. Water Resour. 33, 1524–1541. https://doi.org/10.1016/j.advwatres.2010.09.002

Shen, C., Riley, W.J., Smithgall, K.M., Melack, J.M., Fang, K., 2016. The fan of influence of streams and channel feedbacks to simulated land surface water and carbon dynamics. Water Resour. Res. 52, 880–902. https://doi.org/10.1002/2015WR018086

Stone, P., 2006. Geologic Map of the West Half of the Blythe 30' by 60' Quadrangle, Riverside County, California and La Paz County, Arizona [WWW Document]. USGS Sci. Investig. Map 2922. URL https://pubs.usgs.gov/sim/2006/2922/ (accessed 4.17.17).

Stonestrom, D.A., Constantz, J., Ferré, T.P.A., Leake, S.A., 2007. Ground-Water Recharge in the Arid and Semiarid Southwestern United States [WWW Document]. U.S. Geol. Surv. Prof. Pap. 1703. URL https://pubs.usgs.gov/pp/pp1703/ (accessed 5.26.17).

Stonestrom, D.A., Prudic, D.E., Laczniak, R.J., Akstin, K.C., 2004. Tectonic, climatic, and land-use controls on groundwater recharge in an arid alluvial basin: Amargosa Desert, U.S.A., in: Hogan, J.F., Phillips, F.M., Scanlon, B.R. (Eds.), Groundwater Recharge in a Desert Environment: The Southwestern United States. American Geophysical Union, Washington, DC, pp. 29–47. https://doi.org/10.1029/009WSA03

Sutanudjaja, E.H., van Beek, R., Wanders, N., Wada, Y., Bosmans, J.H.C., Drost, N., van der Ent, R.J., de Graaf, I.E.M., Hoch, J.M., de Jong, K., Karssenberg, D., López López, P., Peßenteiner, S., Schmitz, O., Straatsma, M.W., Vannametee, E., Wisser, D., Bierkens,

M.F.P., 2017. PCR-GLOBWB 2: a 5 arc-minute global hydrological and water resources model. Geosci. Model Dev. Discuss. 1–41. https://doi.org/10.5194/gmd-2017-288

Therrien, R., McLaren, R.G., Sudicky, E.A., Panday, S.M., 2006. HydroGeoSphere: A three-dimensional numerical model describing fully-integrated subsurface and surface flow and solute transport [WWW Document]. Groundw. Simul. Gr. Waterloo Ontario.

Tian, Y., Zheng, Y., Wu, B., Wu, X., Liu, J., Zheng, C., 2015. Modeling surface water-groundwater interaction in arid and semi-arid regions with intensive agriculture. Environ. Model. Softw. 63, 170–184. https://doi.org/10.1016/j.envsoft.2014.10.011

Tonkin, M., Doherty, J., 2009. Calibration-constrained Monte Carlo analysis of highly parameterized models using subspace techniques. Water Resour. Res. 45, 1–17. https://doi.org/10.1029/2007WR006678

USGS, 1995. Groundwater Atlas of the United States, California and Nevada [WWW Document]. United States Geol. Surv. Hydrol. Investig. Atlas HA-730-B. URL http://pubs.usgs.gov/ha/ha730/ch_b/

van Genuchten, M.T., 1980. A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils. Soil Sci. Soc. Am. J. 44, 892. https://doi.org/10.2136/sssaj1980.03615995004400050002x

Wilson, J.L., Guan, H., 2004. Mountain-Block Hydrology and Mountain-Front Recharge, in: Hogan, J.F., Phillips, F.M., Scanlon, B.R. (Eds.), Groundwater Recharge in a Desert Environment: The Southwestern United States, Water Science and Application. American Geophysical Union, Washington, DC. https://doi.org/10.1029/WS009

WorleyParsons, 2010. Genesis Solar Energy Project Groundwater Resources Investigation [WWW Document]. Rep. to Calif. Energy Comm. URL http://www.energy.ca.gov/sitingcases/genesis_solar//documents/applicant/2010-02-11_Groundwater_Resources_Investigation_TN-55916.pdf (accessed 4.18.17).

WorleyParsons, 2009. Groundwater resources investigation, Genesis Solar Energy Project, Riverside Country, California [WWW Document]. Genes. Sol. Energy Proj. Appl. Certif. Vol. I. URL http://www.energy.ca.gov/sitingcases/genesis_solar/documents/applicant/afc/volume_1/ (accessed 8.1.16).

Wösten, J.H.M., Pachepsky, Y.A., Rawls, W.J., 2001. Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. J. Hydrol. 251, 123–150. https://doi.org/10.1016/S0022-1694(01)00464-4

Xu, T., Gómez-Hernández, J.J., 2016. Characterization of non-Gaussian conductivities and porosities with hydraulic heads, solute concentrations, and water temperatures. Water Resour. Res. 52, 6111–6136. https://doi.org/10.1002/2016WR019011

Young, M., McDonald, E., Caldwell, T., Benner, S., Meadows, D., 2004. Hydraulic Properties of a Desert Soil Chronosequence in the Mojave Desert, USA. Vadose Zo. J. 3, 956–963. https://doi.org/10.2113/3.3.956

Zhang, D., Madsen, H., Ridler, M.E., Kidmose, J., Jensen, K.H., Refsgaard, J.C., 2016.

Multivariate hydrological data assimilation of soil moisture and groundwater head. Hydrol. Earth Syst. Sci. 20, 4341–4357. https://doi.org/10.5194/hess-20-4341-2016

Zhang, D., Madsen, H., Ridler, M.E., Refsgaard, J.C., Jensen, K.H., 2015. Impact of uncertainty description on assimilating hydraulic head in the MIKE SHE distributed hydrological model. Adv. Water Resour. 86, 400–413. https://doi.org/10.1016/J.ADVWATRES.2015.07.018

Zovi, F., Camporese, M., Hendricks Franssen, H.-J., Huisman, J.A., Salandin, P., 2017. Identification of high-permeability subsurface structures with multiple point geostatistics and normal score ensemble Kalman filter. J. Hydrol. 548, 208–224. https://doi.org/10.1016/J.JHYDROL.2017.02.056

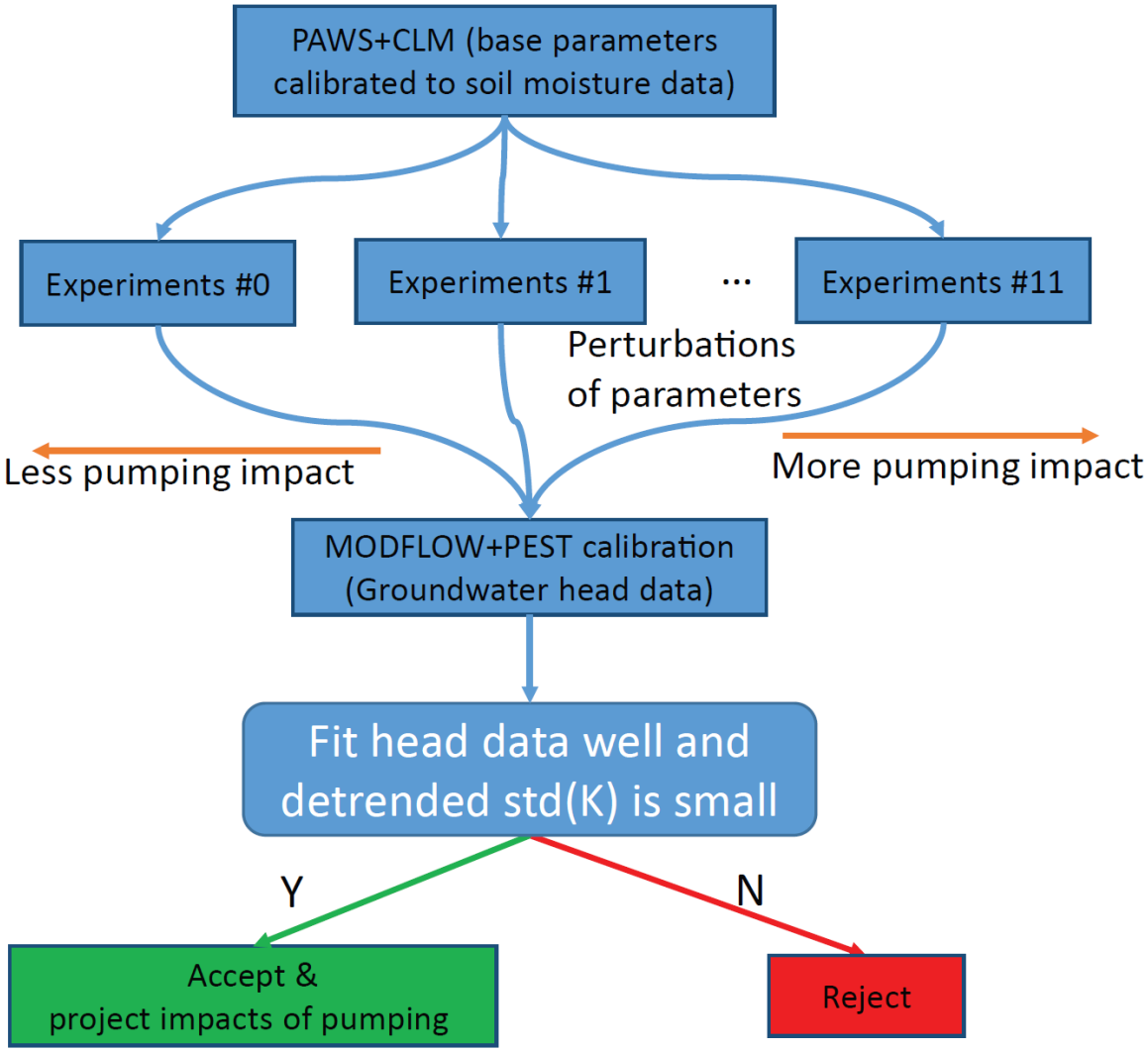


Figure 1. The proposed dual-model approach. We collected 4 years of field soil moisture measurements to estimate base soil properties with a process-based integrated hydrologic model, PAWS+CLM. We then generated a range of recharge estimates by making perturbations to the calibrated soil parameters. Groundwater observations were used to constrain *K* in a parameter estimation package, MODFLOW+PEST and, more importantly, retain or reject some of the recharge estimates. The retained estimates were used to produce the range in possible drawdown induced by solar-plant pumping, given the available information.

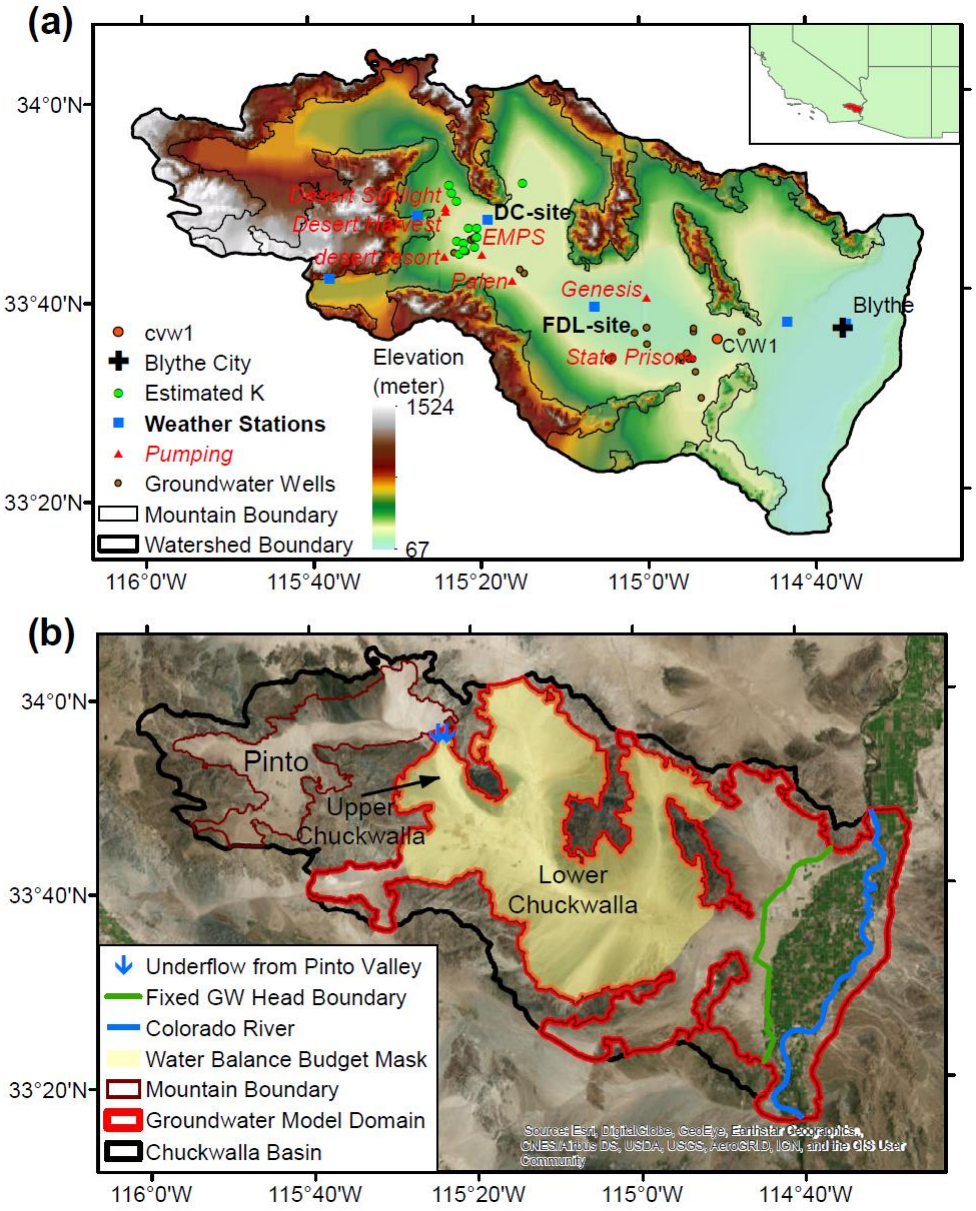
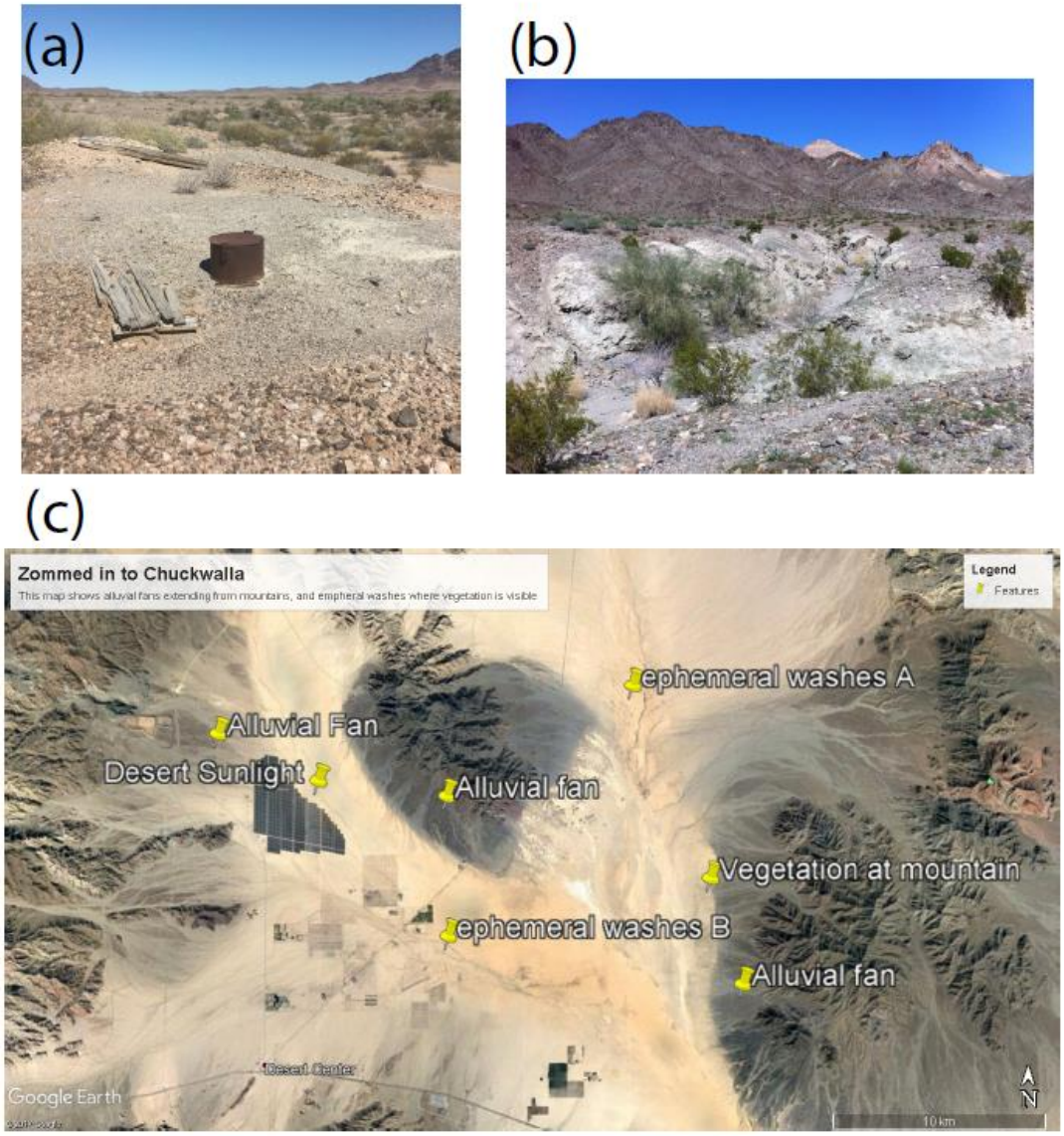


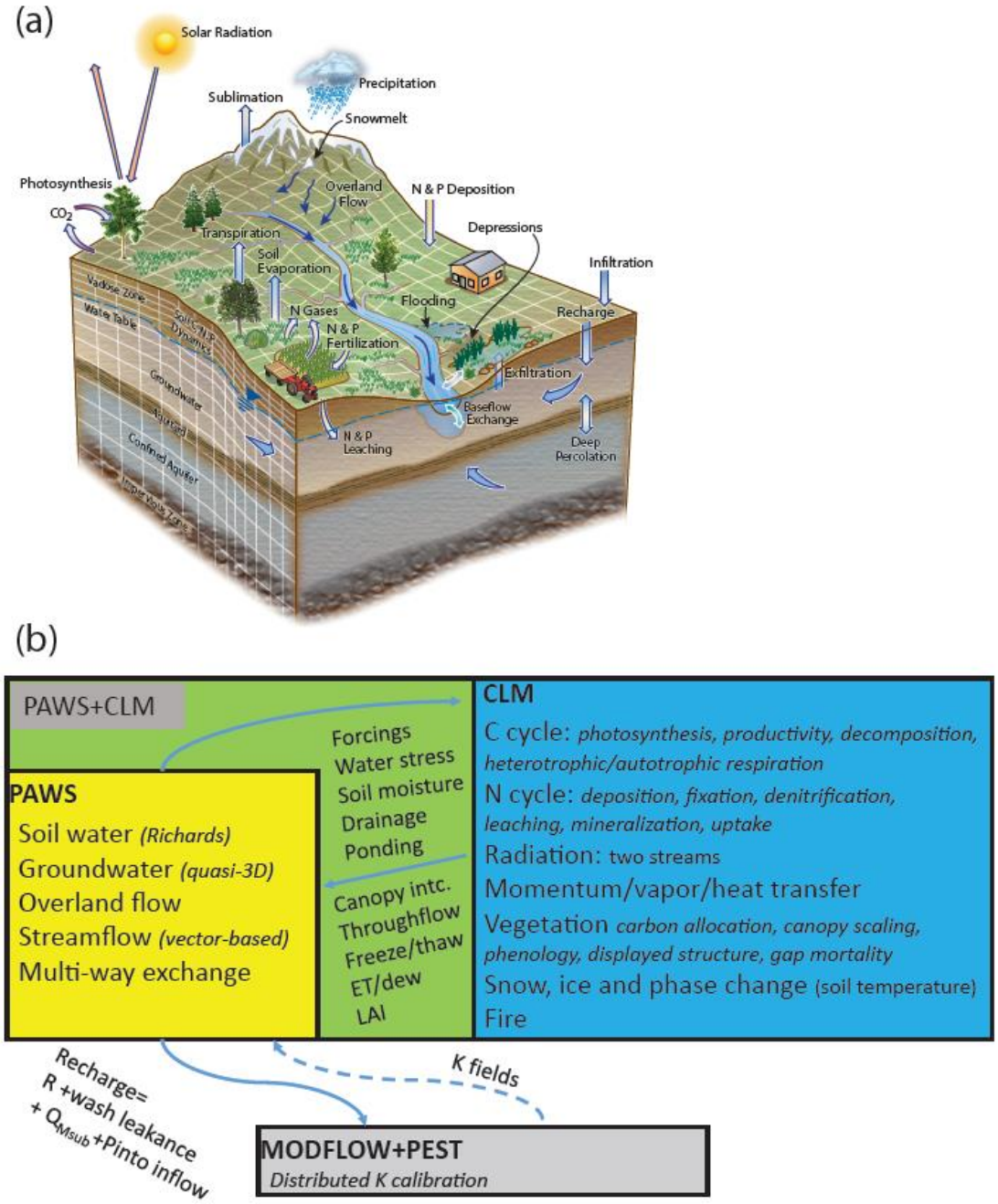
Figure 2. (a) Map showing locations of soil moisture stations including Ford Dry Lake (FDL) and Desert Center (DC), groundwater wells, existing *K* estimates (groundwater observation wells which also had K estimates from pumping tests), existing pumping sources, solar plants, and Eagle Mountain Pumped Storage (EMPS) project. (b) Satellite image of the Chuckwalla basin and the modeling domain. The MODFLOW+PEST model domain is smaller than PAWS+CLM model domain. A fixed head boundary condition (green line), which was constructed by connecting known groundwater head, was set to encompass the agricultural region so that dynamics east of this line do not impact the calibration. The water balance budget mask refers to the area over which mass balance is reported. The MODFLOW model excluded mountains, but mountain-front subsurface inflow were treated as source terms.



923

924 **Figure 3. (a) A well in the basin surrounded by soils with visible desert pavement; (b) the foot of an**
925 **alluvial fan looking upslope to higher elevations. Note that vegetation is visibly denser on the**
926 **alluvial fan. Washes are also visible; (c) a zoomed-in satellite image of the upper Chuckwalla Valley,**
927 **with patterns of ephemeral washes and vegetation annotated.**

928



929

930 **Figure 4. (a) Sketch of PAWS+CLM hydrologic and ecosystem processes (reprinted from (Shen et**
931 **al., 2016) with permission). (b) Information flow chart between PAWS, CLM and**
932 **MODFLOW+PEST.**

933

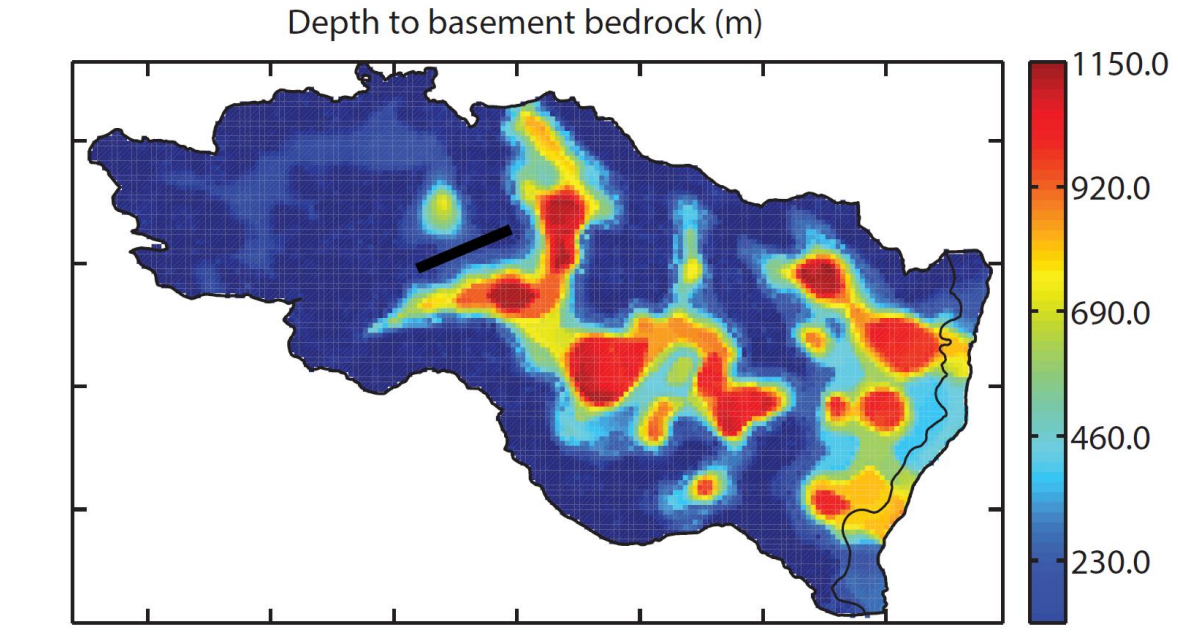


Figure 5. Depth to the basement bedrock map. The thick black line indicates a buried ridge that separates the upper and lower Chuckwalla Valleys that is visible in Figure 6 of Appendix C in (GEI, 2010) **and multiple well-based transect profiles.**

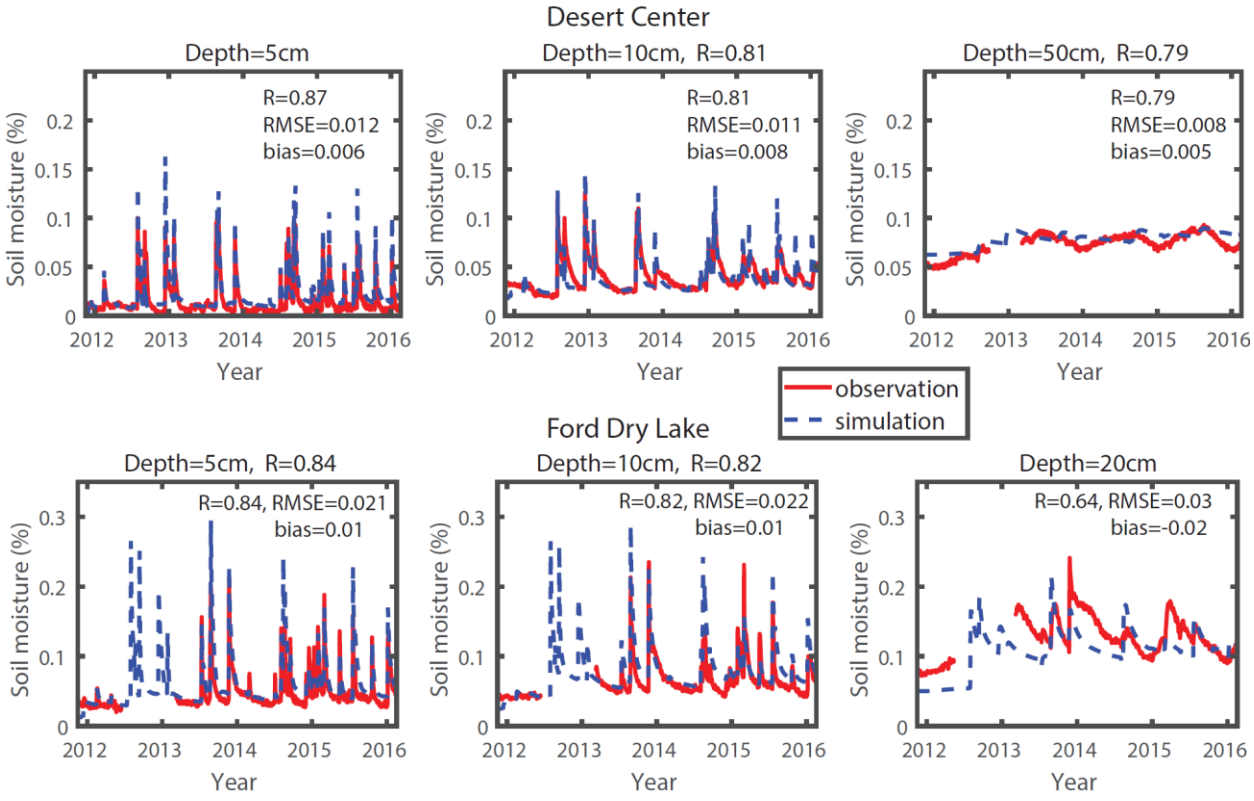


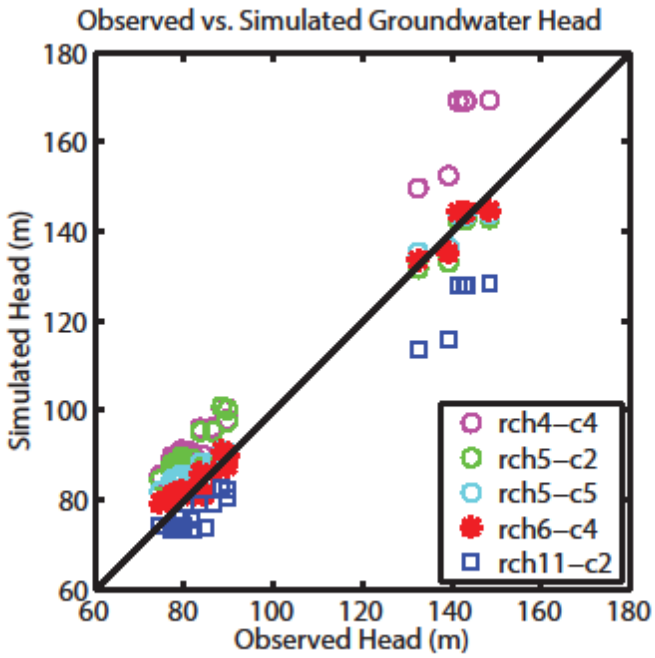
Figure 6. Soil moisture comparisons at the Desert Center site (upper three panels) and the Ford Dry Lake site (lower three panels). *R* is the correlation coefficient between simulation and observation. RMSE is root-mean-squared error. Bias is the difference between the mean.

Recharge	Realizations									
	i		*ii*		*iii*		*iv*		*v*	
1	13.4	14.0	5.6	13.5	11.4	12.8	35.7	36.4	6.9	13.2
	0.0%	6.1	0.0%	6.5	0.0%	7.4	0.0%	2.1	0.0%	8.1
2	86.9	87.1	49.9	51.4	5.9	13.9	45.5	51.2	12.2	13.6
	6.0%	1.3	0.0%	1.3	0.0%	6.6	0.0%	4.1	0.0%	7.8
3	34.6	36.3	4.4	11.0	34.6	36.3	8.9	10.6	8.9	10.6
	0.0%	6.8	0.0%	6.6	0.0%	2.3	0.0%	7.4	0.0%	7.4
4	1.1	11.0	5.2	7.3	1.7	10.4	14.0	15.4	7.1	7.9
	0.0%	6.1	0.0%	6.8	0.0%	6.0	0.6%	6.2	0.3%	6.8
5	11.3	15.3	6.1	8.2	4.5	5.9	6.2	18.9	3.1	4.5
	0.0%	1.3	0.0%	3.6	0.1%	3.7	0.0%	2.2	0.1%	6.8
6	2.0	3.8	1.1	2.7	1.1	2.6	1.1	2.6	1.0	2.6
	0.6%	3.9	20.4%	5.5	2.5%	4.6	8.6%	3.9	4.8%	4.3
7	4.5	5.1	6.7	7.6	1.1	2.6	-0.2	3.5	3.7	4.5
	92.4%	2.0	5.6%	2.2	5.2%	4.3	0.1%	4.5	17.7%	3.6
8	2.8	3.6	0.6	2.3	0.1	4.0	2.6	3.9	-0.3	3.1
	15.3%	2.6	21.7%	3.8	0.0%	5.3	0.1%	5.2	0.2%	5.1
9	0.2	2.2	1.9	3.1	0.4	2.3	1.8	3.2	0.2	2.3
	29.8%	3.9	1.2%	2.2	24.4%	3.4	2.7%	7.9	28.3%	4.1
10	0.1	2.2	0.1	2.2	3.8	5.4	-0.5	2.5	1.7	5.4
	13.4%	3.9	18.9%	2.1	13.3%	3.8	1.0%	4.0	2.4%	6.3
11	-11.5	14.6	-9.0	10.8	-11.7	15.1	-8.2	9.8	-1.6	4.5
	0.0%	4.1	0.0%	4.6	0.0%	5.7	0.1%	5.5	14.7%	3.8
12	-6.5	7.4	-11.1	13.3	-15.5	19.8	-6.4	7.0	-14.7	18.9
	15.2%	5.0	0.1%	3.5	0.0%	7.8	93.0%	4.4	0.0	6.9

Legend

bias	rmse
P_E	σ

Figure 7. Detailed metrics for model rejection. Green-filled cases pass all statistical tests. A colored box indicates a model rejection based on the criteria the legend as described in Section 2.5. As shown in the legend, for every calibrated field (12 recharge values, each with 5 calibration realizations), the numbers shown for each field are mean bias (upper left) of residuals (calibrated-observed head), root-mean-squared error (rmse, upper right), *p-value* for the elevation regression test (P_E, lower left), and standard deviation of the detrended *K* residuals (σ_K, lower right).

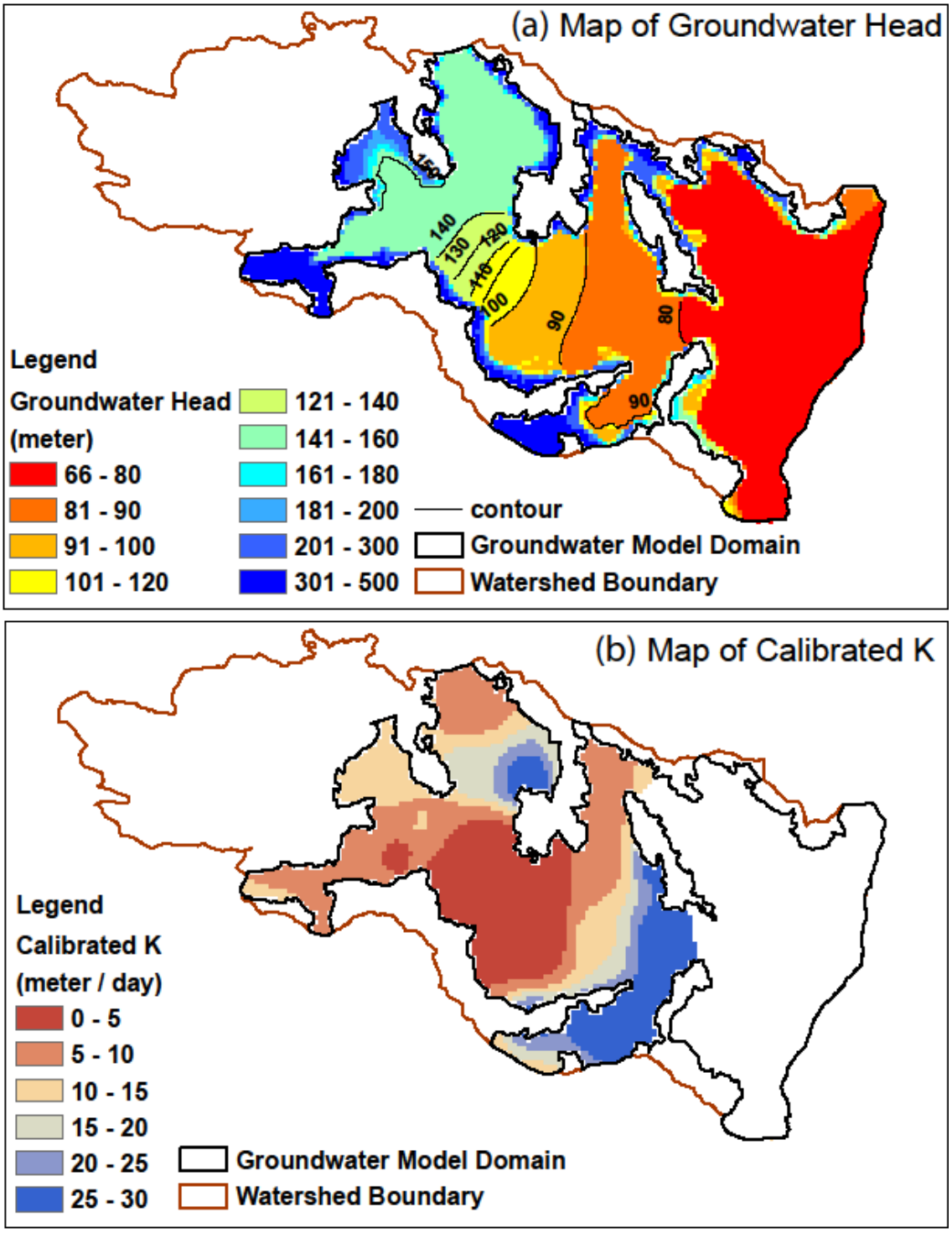


953

954 **Figure 8. Observed vs. calibrated groundwater head for accepted recharge rates, those that lead to**
955 **over-estimation of head, and under-estimation of head. A filled-in symbol indicates acceptance.**
956 **"rch6-c3" means the calibration realization 3 (with a particular initial guess for *K*) using recharge**
957 **from simulation #6. Other data series are defined similarly.**

958

959



960

961 **Figure 9. (a) Simulated groundwater head map and (b) corresponding calibrated groundwater**
962 **conductivity from simulation #6, the base case. This map does not include the effects of the assumed**
963 **solar plant pumping. The Palo Verde Mesa Valley groundwater basin near Blythe (to the East of**
964 **the mountain mouth) is controlled by the fixed head boundary condition.**

965

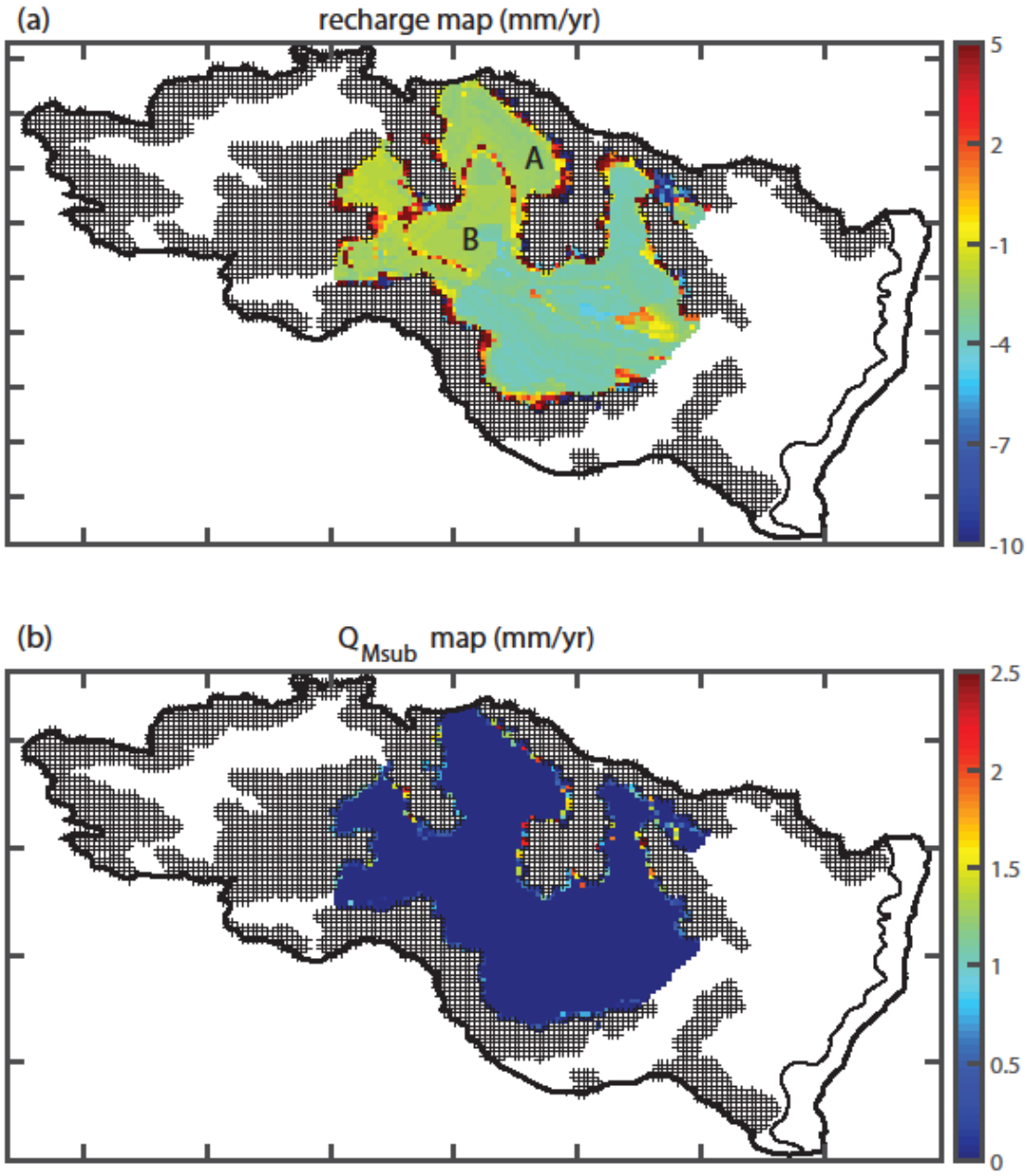


Figure 10. Simulated recharge maps from the base parameter set (simulation #6) (a) The total recharge, consisting of run-on percolation, soil matrix recharge, and mountain front subsurface recharge (Q_{MSub}).; (b) mountain-front subsurface recharge, which is lateral subsurface flow from thin mountain soils. The cross hatched areas are the bedrock / mountain exposures. The Palo Verde Mesa Basin / Colorado River Floodplain (white area in the east) are not considered in the calibration. The Pinto Valley (white area to the northwest) is outside of the groundwater modeling domain, however, groundwater inflow to the Chuckwalla Basin was included as an input term.

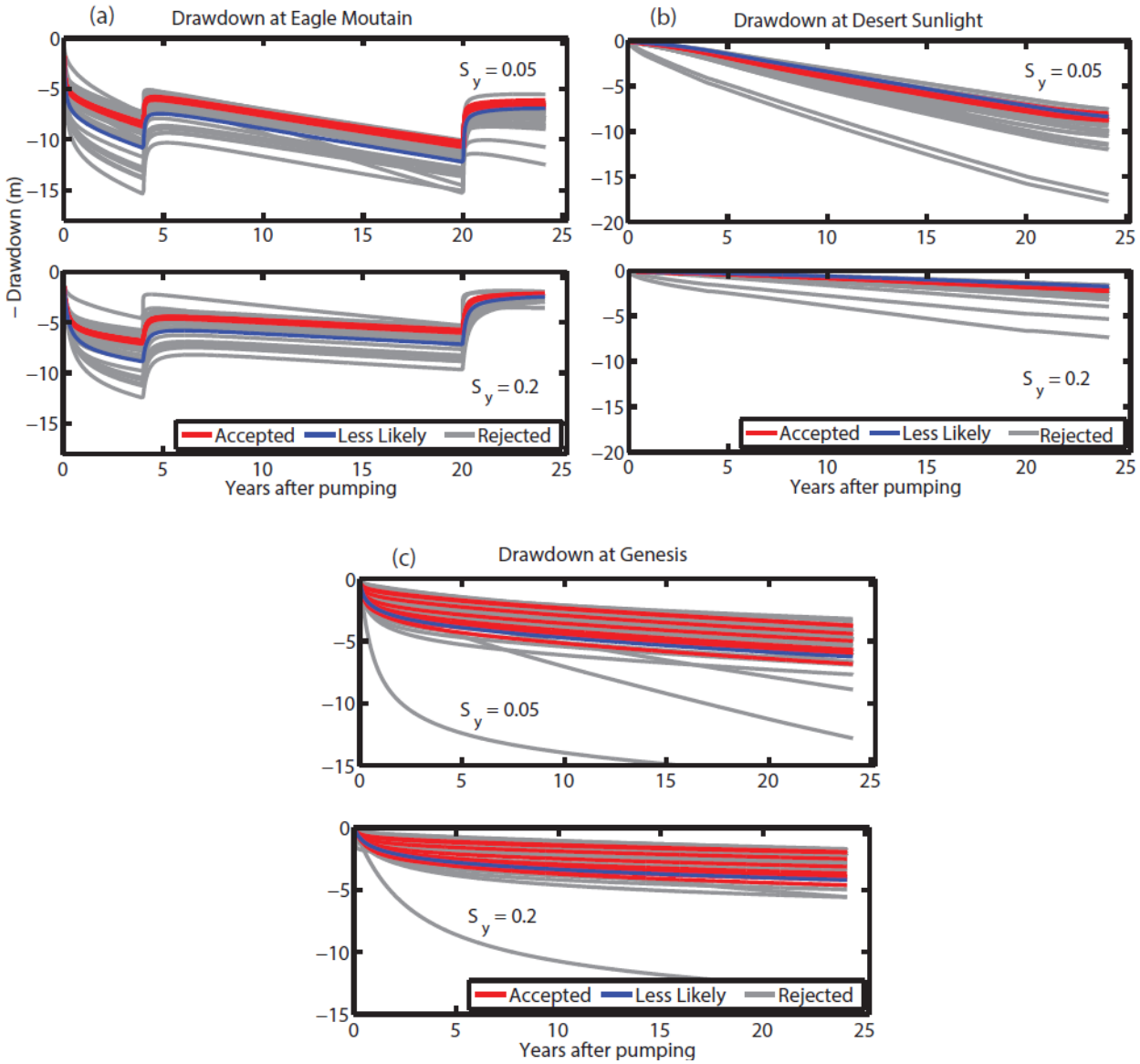


Figure 11 (a) Influence of assumed pumping on the water table at the Eagle Mountain Pumped Storage (EMPS) Project pumping site. The first four years is the initial fill phase. 5-20 years is the re-fill period. The pumping was terminated after 20 years to examine the rate of recovery. The red lines indicate accepted recharge rates. The magenta lines are "less-likely" recharge rates that have higher error statistically but could not be completely rejected. The gray lines are the rejected recharge rates indicating the extent of uncertainty facing the prediction if no model rejection was applied. (b) Same figure as in (a) but for the Desert Sunlight solar plant; (c) the same Figure as (a) but for the Genesis solar plant.

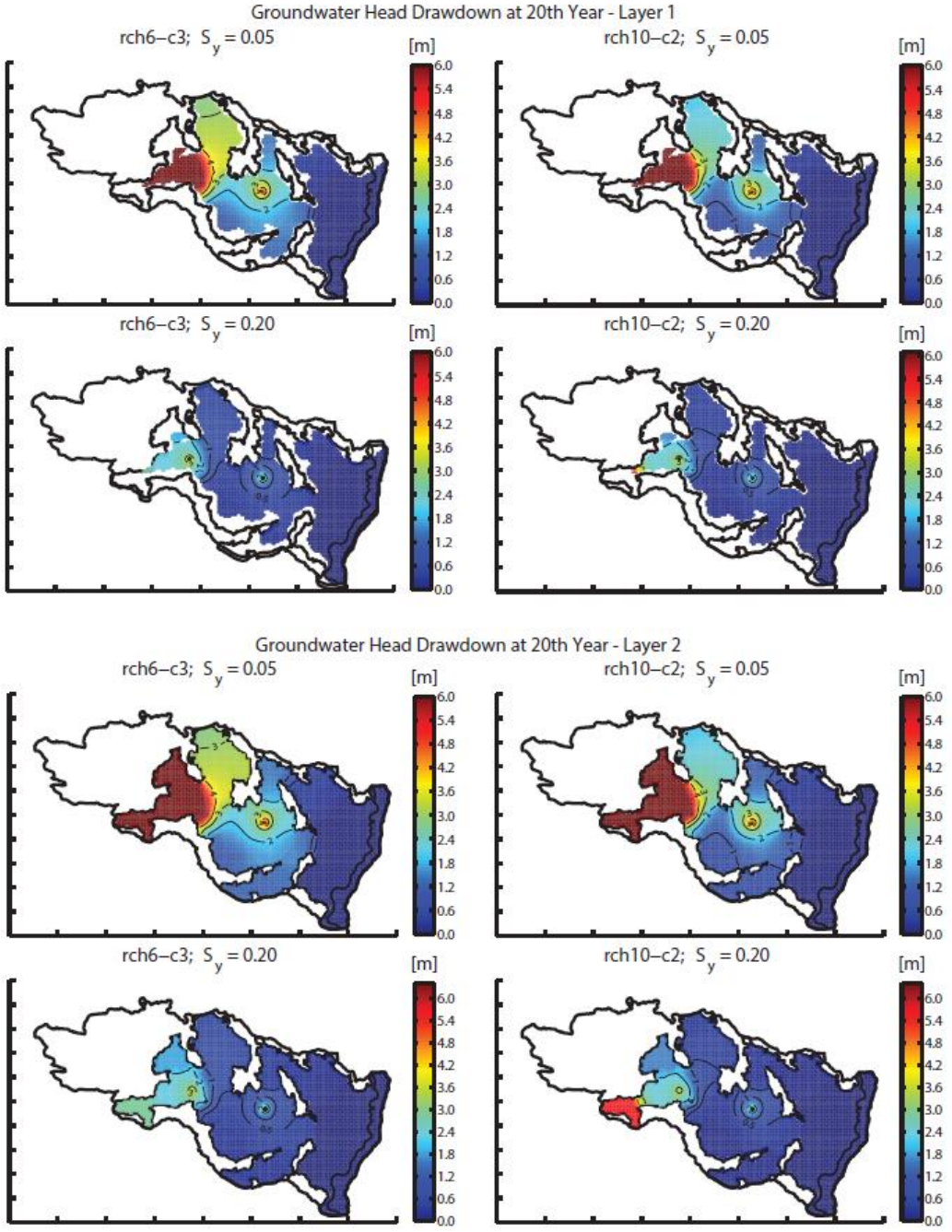


Figure 12. The cones of depression formed by drawdown (groundwater head from simulations without pumping minus that with pumping) for different assumptions for S_y.

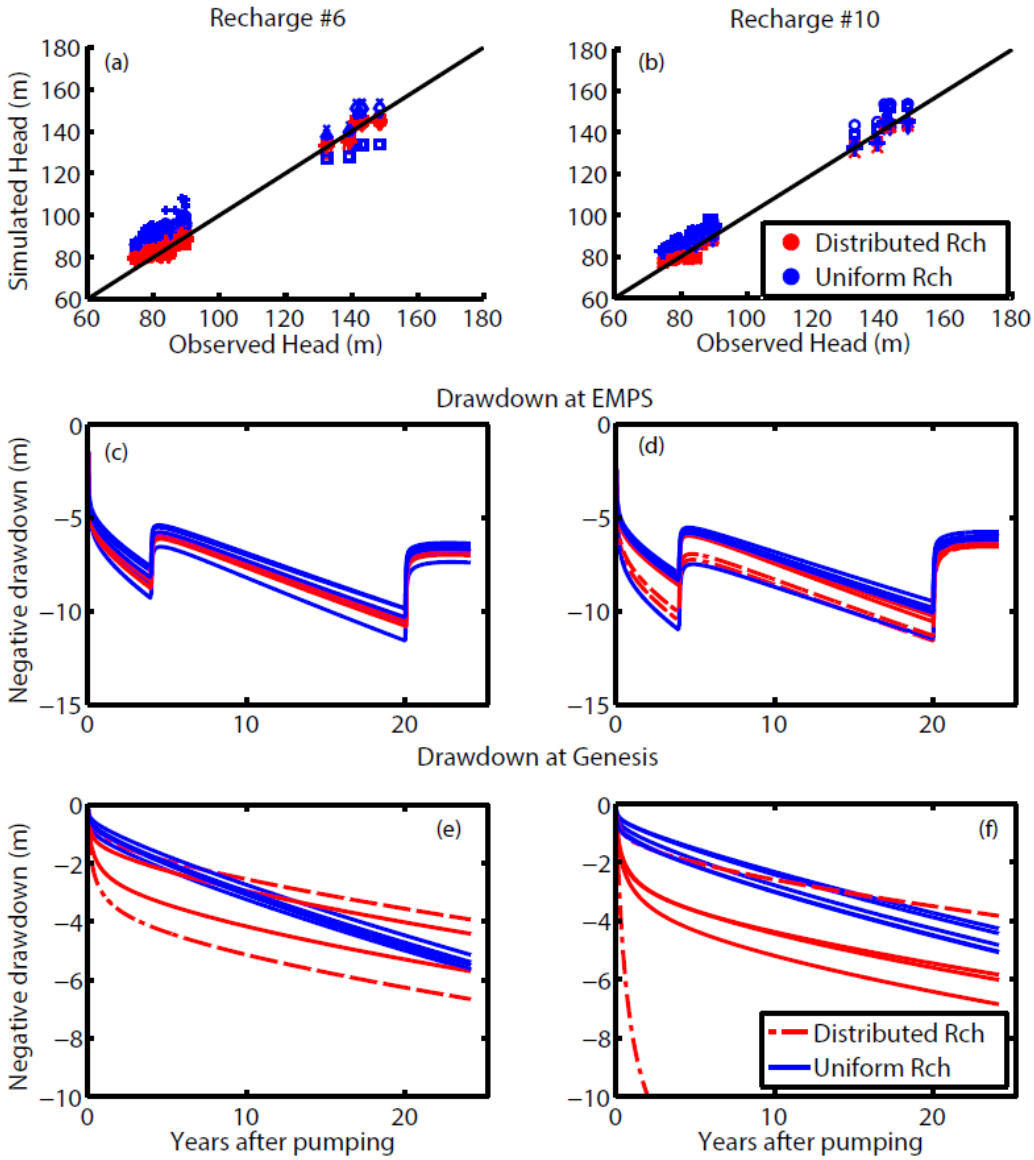


Figure 13. Comparing model-estimated recharge vs. uniform recharge for $S_y = 0.15$. No model rejection was applied to uniform-recharge simulations. Dashed lines indicates distributed-recharge simulations that have been rejected. (a-b) calibrated vs observed groundwater head for simulations #6 and #10. Different symbols indicate different calibration cases; (c-f) projected impacts of pumping at EMPS and Genesis.

Table 1. Modeled pumping sources in the Chuckwalla Basin Valley. We excluded sources to the east of the groundwater basin boundary. Superscripts: [o] operational and [p] proposed or planned.

	Pumping Rate (10^3 m^3/yr)	Reference
Genesis [o]	1881	(WorleyParsons, 2009)
Desert Sunlight [o]	64	(BLM, 2011)
Desert Harvest [o]	65	(EPA, 2012)
Palen [p]	271	(BLM, 2017)
EMPS (1-4 yrs)	9992	(GEI, 2010)
EMPS (5-20 yrs)	2220	
EMPS (21-24 yrs)	0	
State prison [o]	2590	(WorleyParsons, 2009)
Desert Resort [o]	1344	(WorleyParsons, 2009)
Existing pumping (prison+resort+misc.)	3935	
Total solar—initial-fill (period 1, 1-4 yrs)	16209	
Total solar—re-supply (period 2, 5-20 yrs)	8438	
Total solar—Decommissioned (period 3, 21-24 yrs)	6217	
20-year annualized from all solar plants	9992	

Table 2. Calibrated soil parameters for two field sites. The van-Genuchten parameters K_S, N, α were kept constant throughout different depths. θ_r and λ were adjusted at different depths to better fit the data.

Ford Dry Lake						
Depth	***Ks* (m/day)**	***N* (-)**	**α (m^{-1})**	**θ_r (-)**	**θ_s (-)**	**λ (-)**
5 cm	0.1	1.6	4	0.00	0.3805	-1.2155
10 cm	0.1	1.6	4	0.00	0.4221	-0.1059
20 cm	0.1	1.6	4	0.02	0.4221	-0.1059
50 cm	0.1	1.6	4	0.05	0.4221	-0.1059
Desert Center						
Depth	***Ks* (m/day)**	***N* (-)**	**α (m^{-1})**	**θ_r (-)**	**θ_s (-)**	**λ (-)**
5 cm	0.12	1.8	3.2	1.00E-10	0.3877	-1.3
10 cm	0.12	1.8	3.2	1.00E-10	0.3824	-1.3
20 cm	0.12	1.8	3.2	0.025	0.3969	-1
50 cm	0.12	1.8	3.2	0.06	0.3969	-0.8

Table 3. Parameter perturbations for the numerical simulations. These changes were applied as multipliers or additions to default values (Table 2). N/C means no change is applied. Going from Sim #1 to Sim #11, the resulting recharge decreased. *K* mostly influences Pinto Valley Basin underflow. Simulation #12 was derived from #11: it used the same spatial distribution of recharge but multiplied the values by 0.8.

Parameter	K_S	α	K	K_S for mountain areas	Deep layer porosity for non-mountain areas	N
sim#1	× 10	× 1.5	× 3	N/C	× 1.2	N/C
sim#2	× 8	× 1.4	× 2.5	N/C	× 1	N/C
sim#3	× 6	× 1.3	× 2	N/C	× 1	N/C
sim#4	× 4	× 1.2	× 1.5	N/C	× 1	N/C
sim#5	× 2	× 1.1	× 1.25	N/C	× 1	N/C
sim#6	× 1	× 1	× 1	N/C	× 1	N/C
sim#7	× 1	× 1	× 1	=1.6 m/day	× 0.8	N/C
sim#8	× 0.75	× 0.85	× 0.5	N/C	× 0.7	N/C
sim#9	× 0.5	× 0.7	× 0.3	N/C	× 0.55	N/C
sim#10	× 0.5	× 0.7	× 0.3	=1.6	× 0.45	N/C
sim#11	× 0.5	× 0.7	× 0.3	N/C	× 0.55	-0.2
sim#12	× 0.5	× 0.7	× 0.3	N/C	× 0.55	-0.2

Table 4. Mass balance (10^3 m^3/yr) and model acceptance status from the perturbed simulations. These fluxes were summed up for the "water balance budget mask" area in Figure 2b. 'Q_{Msub}' means mountain-front subsurface recharge. A recharge was rejected when none of the 5 realizations was retained.

Recharge #	Soil & wash recharge	Pinto underflow	Q_{Msub}	Total inflow	Prcp	Annualized pumping	Results
sim#1	22830	2758	368	25956	253327	9992	***Reject – always overestimate head***
sim#2	22898	2880	390	26168	253327	9992	***Reject – always overestimate head***
sim#3	20856	2192	297	23345	253327	9992	***Reject – always overestimate head***
sim#4	18565	1495	275	20335	253327	9992	***Reject – always overestimate head***
sim#5	15719	1248	278	17244	253327	9992	***Reject—either GW is over-estimated or K variation is too large***
sim#6	12924	1082	259	14264	253327	9992	Accept 2 runs
sim#7	13067	1018	224	14311	253327	9992	Accept 1 run
sim#8	11702	644	213	12561	253327	9992	Accept 1 run
sim#9	10533	459	168	11159	253327	9992	Accept 3 runs
sim#10	9743	479	132	10354	253327	9992	Accept 2 runs
sim#11	6549	395	236	7179	253327	9992	*Mostly rejected. One narrow retention retained as "unlikely"*
sim#12	5239	395	236	5869	253327	9992	***Reject – always underestimate GW head***